

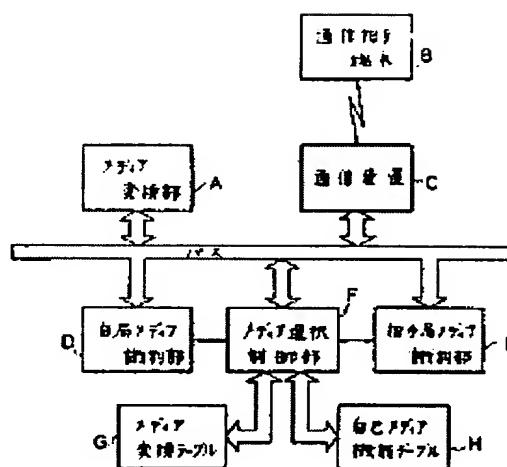
MEIDA CONVERTING SYSTEM

Patent number: JP63086655
Publication date: 1988-04-18
Inventor: YOSHIMURA SUSUMU; others: 04
Applicant: TOSHIBA CORP
Classification:
 - **International:** H04M11/00; G06F13/00; H04L13/00
 - **European:**
Application number: JP19860230061 19860930
Priority number(s):

Abstract of JP63086655

PURPOSE: To communicate information with media which can be treated by respective terminal in various forms of terminals by detecting the input media of the information which is going to be communicated and the media with which the opposite terminal of communication can process and selecting the form of the conversion of media according to the detected results.

CONSTITUTION: A self-station media identification part D detects the input media of the information which should be communicated and at the same time an opposite station media identification part E detects the media with which the opposite terminal of communication can process to obtain the communication media for communicating the information. Under the control of a media selection control part F, the kind of the conversion of media which is necessary to communicate the input information to the opposite terminal of communication is obtained by referring to a media conversion table G based on the input media and the communication media, and media conversion means A is selectively started to transmit the input information to the opposite terminal of communication after the conversion of media according to the kind of the conversion of media. Thus the information which should be communicated can be communicated by converting to the media according to the kind of media with which the opposite terminal of communication can process.



Data supplied from the **esp@cenet** database - Worldwide

⑪ 公開特許公報 (A)

昭63-86655

⑤Int.Cl.*

H 04 M 11/00
G 06 F 13/00

識別記号

302
351
354

庁内整理番号

8020-5K
R-7218-5B
7218-5B

⑩公開 昭和63年(1988)4月18日

※審査請求 未請求 発明の数 1 (全36頁)

⑪発明の名称 メディア変換方式

⑪特願 昭61-230061

⑪出願 昭61(1986)9月30日

⑪発明者 吉村 晋 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝総合研究所内

⑪発明者 鈴木 謙二 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝総合研究所内

⑪発明者 栗原 基 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝総合研究所内

⑪発明者 斎藤 裕美 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝総合研究所内

⑪出願人 株式会社東芝 神奈川県川崎市幸区堀川町72番地

⑪代理人 弁理士 鈴江 武彦

最終頁に続く

明細書

1. 発明の名称

メディア変換方式

2. 特許請求の範囲

(1) 複数種類のメディア変換手段を備えたマルチメディア対応のワークステーションであって、

該ワークステーションは、通信に供される情報の入力メディアを検出する入力メディア検出部と、通信相手端末が処理可能なメディアを検出して上記情報を通信する為の通信メディアを求める相手メディア検出部と、メディア変換テーブルを参照して上記入力メディアと通信メディアに応じてメディア変換の種別を求める手段とを具備し、

求められたメディア変換の種別に応じて前記メディア変換手段を選択的に起動して前記入力情報をメディア変換して前記通信相手端末に送信してなることを特徴とするメディア変換方式。

(2) メディア変換手段は、音声認識装置、文字認識装置、图形認識装置、イメージ認識装置、音声合成装置、イメージ合成装置、图形合成装置や、

各種データの圧縮・伸長装置等からなり、メディア変換の種別に応じて選択的に用いられるものである特許請求の範囲第1項記載のメディア変換方式。

(3) 相手メディア判定部は、通信相手端末の種別を検知して該通信相手端末が処理可能なメディアの種別を求めるものである特許請求の範囲第1項記載のメディア変換方式。

(4) 相手メディア判定部は、通信相手端末が処理可能なメディアが複数種類あるとき、その1つを選択して通信メディアを定めるものである特許請求の範囲第1項記載のメディア変換方式。

(5) 相手メディア判定部は、通信相手端末が通信メディアに応じて処理メディアを自動設定するか否かを判定して情報を通信するに適した通信メディアを求めるものである特許請求の範囲第1項記載のメディア変換方式。

(6) 相手メディア判定部は、通信相手端末の要求に応じて該ワークステーションの種別、または該ワークステーションが出力可能なメディアの種別

を該通信相手端末に通知する機能を備えてなるものである特許請求の範囲第1項記載のメディア変換方式。

3. 発明の詳細な説明

【発明の目的】

(産業上の利用分野)

本発明は、マルチメディア対応のワークステーションにおけるメディア変換方式に関する。

(従来の技術)

情報処理技術の発展に伴い、種々のワークステーションが開発されている。そして従来一般的なコード情報のみならず、音声データや画像データ等の種々のメディアの情報を処理し得るワークステーションが開発されている。またその通信メディアも多様化している。

ところでこの種のマルチメディア対応のワークステーションは、専らオフィス毎に設置され、そのオフィス内でのみ使用されているのが実情である。また最近では複数のワークステーションを通じて相互に結び、複数のワークステーション

から通信相手端末に通信すべき情報を、該通信相手端末が処理可能なメディアの種類に応じたメディアに変換して情報通信することを可能ならしめ、ワークステーションが持つ情報を種々の情報端末にて出力できるようにしたメディア変換方式を提供することにある。

【発明の構成】

(問題点を解決するための手段)

本発明は第1図にその概略を示すように、音声認識装置、文字認識装置、图形認識装置、イメージ認識装置、音声合成装置、イメージ合成装置、图形合成装置や、各種データの圧縮・伸長装置等の複数種類のメディア変換部Aと、種々の通信相手端末との間で情報通信を行なう通信装置Cとを具備したマルチメディア対応のワークステーションであって、

通信に供される情報の入力メディアを入力メディア検出部(自局メディア識別部D)にて検出すると共に、相手メディア検出部(相手局メディア識別部E)にて通信相手端末が処理可能なメディ

アを検出して上記情報を通信する為の通信メディアを求め、メディア選択制御部Fの制御の下で上記入力メディアと通信メディアに従ってメディア変換テーブルGを参照して前記入力情報を前記通信相手端末に通信する為に必要なメディア変換の種別を求め、このメディア変換の種別に応じて前記メディア変換手段Hを選択的に起動して前記入力情報をメディア変換して前記通信相手端末に送信するようにしたことを特徴とするものである。

具体的には、例えば通信相手端末がファクシミリ装置を備えた電話端末である場合、その通信相手端末の種別を検出して画像データや音声データを通信可能なことを求め、この通信データのメディアに応じて、例えば文字コード列で示される文書情報を文字画像情報にメディア変換して情報通信するようにしたものである。

またメディア選択制御部Fは必要に応じて自己メディア機能テーブルHを参照して該ワークステーションが備えたメディア変換機能等を求める、上記メディア変換の種別の選択を制御するものとな

っている。

(作用)

本発明によれば、通信相手端末がワークステーションと同じ処理機能を備えていなくても、ワークステーションの相手メディア検出部がその通信相手端末が処理できるメディアの種別を、例えば通信相手端末から送られてくるメディア識別信号から検出し、この検出結果に従って上記通信相手端末に情報通信するに適したメディアの種別を求めてメディア交換手段を選択的に起動し、前記入力情報のメディアを通信相手端末が処理可能なメディアに変換して通信する。従って通信相手端末が備えている処理機能(処理メディア)に応じてワークステーションが持つ情報のメディアを交換し、これを効果的に通信出力することが可能となる。

例えばワークステーションに接続された通信相手端末がファクシミリ装置の場合、文字コード列で示される文書情報を文字画像として情報通信して出力することが可能となり、また電話端末が接

る。

位置座標入力装置4: タブレットやマウス等からなり、指定された位置座標情報を入力する。

音声入力部5: マイクロフォン等により構成され、音声情報を入力する。

キーボード部6: 鍵盤のキーを備え、文字・記号コードや制御コード等を入力するものである。

I Cカード部7: 後述するようにI Cカードが装着され、該I Cカードとの間で必要な情報を入出力するものである。

バスコントローラ8: バス1を介する各部間の情報転送を制御する。

音声出力部9: スピーカ等からなり、音声情報を出力する。

ディスプレイ部10: CRTディスプレイや液晶ディスプレイ等からなり、文字・図形・画像等を表示する。

イメージ出力装置11: FAXやカラープリンタ等からなり、種々のイメージ情報をプリント出力

される場合には文書情報を合成音声として通信出力することが可能となる。

またワークステーションでは、例えば文字画像として入力された情報を文字認識して文字コード列化し、更にこの文字コード列の情報を音声情報に変換して電話端末に通信出力することが可能となる。

(実施例)

以下、図面を参照して本発明の実施例について説明する。

第2図は本発明の実施例に係る知的ワークステーションの概略構成図である。この知的ワークステーションは、以下の各部を備えて構成される。

バス1: 以下に説明する各部の間での必要な情報転送を行なうのに用いられる。

制御部2: マイクロプロセッサを主体として構成され、該知的ワークステーションの各部の動作をそれぞれ制御するものである。

イメージ入力装置3: カメラやスキャナ、OCR等からなり、各種のイメージ情報を入力す

する。

通信装置12, 13: 該ワークステーションと電話機、或いは遠隔地に設置された他のワークステーションや端末等との情報通信を行なう。

切換え装置14: 複数の通信装置を切換え使用する。

タイマー部15: 該ワークステーションに時刻情報や時間情報を提供する。

暗号化処理部16: 種々の情報を暗号化処理する。

音声処理部17: 与えられた音声情報が特定の音声であるか否かを処理する。

イメージ処理部18: 与えられたイメージ情報を特定のイメージであるか否かを処理する。

音声認識部19: 与えられた音声情報を認識処理する。

音声分析部20: 音声入力部5等から入力された音声の特徴を抽出する等して該音声を分析処理する。

文字認識部21: 前記イメージ入力装置3等から入力された文字・記号パターンを認識処理する。

イメージ認識部23：前記イメージ入力装置等から入力された图形イメージ等を認識処理する。

出力形態選択部24：該ワークステーションから出力する情報の形態を選択制御する。

作業環境データ収集部25：該ワークステーションの機能状態や、それによるオフィス内の作業環境等の情報を収集入力する。

音声合成部26：処理データに従って合成音声を生成する。

イメージ合成部27：複数のイメージ情報を合成処理したり、処理データに従ってイメージの編集処理を実行する。

图形合成処理部28：種々の图形を合成処理したり、処理データに従って图形の加入・削除等の編集処理を実行する。

音声の圧縮・伸長部29：音声データを圧縮符号化したり、圧縮された音声データの復元伸長を行なう。

イメージの圧縮・伸長部30：イメージ・データを圧縮符号化したり、圧縮されたイメージ・データ

ードの一端部に、上述した構成の知的ワークステーション本体に接続する為のインターフェース部7b、および表示窓部7cを設けて構成される。

尚、表示窓部7cは透明偏光体を埋め込んで形成されるもので、その位置はインターフェース部7bや半導体回路と乗せしない位置に設定される。またカード本体7aは、上記表示窓部7cに対応する部分のみが透明であっても良く、またその基板全体が透明なものであっても良い。

しかしてICカードは、具体的には第4図にその分解斜視図を示すように、一対のカバー基板7d、7e、これらのカバー基板7d、7eに挟持される埋め込み基板7f、コアシート材7g、プリント基板7hを一体的に熱圧着して構成される。

このプリント基板7hの前記インターフェース部7bに対向する位置には入出力端子7iが設けられ、また表示窓部7cに対向する位置には液晶表示装置7jが設けられる。更にはプリント基板7hには半導体集積回路7kが設けられる。またカバー基板7eには前記プリント基板7hにおける発熱を発散する為

タの復元伸長を行なう。

信号処理部31：種々の信号情報の符号化圧縮やその復元伸長、必要な情報の付加等の一連の信号処理を実行する。

データベース部32：種々の情報を複数のリレーションにそれぞれ分類し、データベースとして蓄積する。尚、このデータベースはコード情報のみならず、イメージや音声等としても構築される。

本発明に係る知的ワークステーションは、基本的には上述した各部を備えて構成され、上述した各部がそれぞれが持つ機能を有効に利用して全体的にインテリジェンスな機能を有するものとなっている。

次に前述したキーボード部5等のように一般的ではなく、この知的ワークステーションにおいて特徴的な機能を有するICカード部7や暗号化処理部16等について更に詳しく説明する。

先ずICカードは、例えば第3図に示すように名刺大の大きさのカード本体7a内にマイクロプロセッサやメモリ回路等の半導体回路を内蔵し、カ

の金属箔7aが設けられる。

尚、カバー基板7d、7eや埋め込み基板7f、コアシート材7gにそれぞれ穿たれた孔部はプリント基板7hに集成された半導体集積回路7k等にそれぞれ対向する位置に設けられたものである。これらの孔部に上記半導体集積回路7k等を嵌合させて前記カバー基板7d、7e、埋め込み基板7f、コアシート材7g、プリント基板7hが積層一体化されてICカードが構成される。そして入出力端子7iは、カバー基板7dに穿たれた孔部を介して露出し、ワークステーション本体に電気的に接続されるインターフェース部7bを構成する。

尚、前記液晶表示装置7jは、例えば第5図にプリント基板7k部の断面構造を示すように、スペーサを介して設けられた一対のポリエーテルサルファンフィルム基板の間に液晶層を挟持し、該フィルム基板の内側面に透明導電膜をそれぞれ形成すると共に、下面側のフィルム基板に偏光体や反射体を設けて構成される。このようにポリエーテルサルファンフィルム基板を用いて液晶表示装置7j

を構成すれば、その厚みを 0.8mm 以下にすることも容易であり、ガラス基板を用いて液晶表示装置を構成する場合に比較して I C カード自体を薄くすることができる。

またこの I C カードの駆動電源については、前記インターフェース部 7b を介してワークステーション本体側から供給するようにしても良いが、カード内に内蔵するようにしても良い。この場合には、例えば高分子フィルムを用いたシート状の電池として組込むようにすれば良い。

しかして前記半導体集積回路 7k は、例えば第 6 図に示すように C P U 7p やデータメモリである P R O M 7q、E 2 P R O M 7r、およびこれらのメモリに対する選択部 7s 等を備えて構成される。P R O M 7q は消去・再書き不可能な大容量の不揮発性メモリであり、前記 C P U 7p に対する制御プログラムや、永久記録すべき情報等を格納している。また E 2 P R O M 7r は書き換え可能な小容量の不揮発性メモリであり、例えば情報の取引番号や、情報通番等の使用時に更新される情報が格納され

そして第 8 図にその概念を示すように、与えられた通信原文を暗号鍵に従って暗号化してその暗号通信文を生成したり、また逆に与えられた暗号通信文を暗号鍵に従って復号してその原文を求める処理を実行する。

秘密鍵ファイル部 16c および公開鍵ファイル部 16d はこの暗号・復号化に用いられる鍵を記憶するものであり、鍵更新部 16e はこれらのファイルされた鍵の更新を司る。

ここで秘密鍵は、この暗号化処理部 16 を所有するワークステーションのみが知る鍵であり、他のワークステーション等に対しては秘密にされる。これに対して公開鍵は各ワークステーションに設定された各秘密鍵とそれぞれ対をなすものであり、他のワークステーションにそれぞれ与えられて公開される。公開鍵ファイル部 16d は、これらの複数のワークステーションがそれぞれ公開した公開鍵を、各ワークステーションに対応して記憶するものである。

暗号化部 16a は第 9 図に示すように、R S A 处

理部 16f と暗号化種別付加部 16j とを備えて構成される。そして通信原文を暗号化して情報通信しようとするとき、その通信相手先のワークステーションが公開した公開鍵を用いて通信原文を暗号化し、その暗号通信文に暗号の種別を示す情報を付加して通信情報を作成し、これを通信するものとなっている。尚、暗号の種別の情報は、例えば “0” で暗号化していないこと、また “1” で暗号化していることを示す情報や、暗号方式を示す情報等からなる。

また復号化部 16b は、自己ワークステーションが公開した公開鍵を用いて成るワークステーションが暗号化して通信してきた暗号通信文を入力し、これを該秘密鍵に対応した秘密鍵を用いて復号化するものであり、第 10 図に示すように暗号文分割部 16k、暗号種別判定部 16l、切換部 16m、16n、R S A 処理部 16q を備えて構成される。

暗号文分割部 16k は、前述したフォーマットで通信されてきた通信情報を前述した暗号種別の情報と暗号化通信文とに分割するものであり、暗号

種別判定部16aは該暗号種別情報からその通信文が暗号化されているか否かを判別している。そして暗号化されていない場合にはその通信文を切換え部16n, 16pを介して出力し、暗号化されている場合にはその通信文をRSA処理部16qに導いている。このRSA処理部16qにて前記秘密鍵を用いて暗号化通信文が復号化処理され、切換え部16pを介して出力される。

尚、RSA処理部16l, 16qは、例えば第11図に示すようにブロック分割部16sとべき乗・剰余計算部16l、およびブロック連結部16uとを組立て構成される。

ここでブロック分割部16sは与えられた信号系列を一定の長さのブロックM₁に分割するものであり、べき乗・剰余計算部16lは各ブロックM₁毎に暗号化の鍵kを用いて

$$N_1 = M_1^k \pmod{n}$$

なる信号系列N₁を求めている。但し、nは固定の値である。この信号系列N₁がブロック連結部16uを介して順に連結されて出力される。

が復号し得るものとなっている。

尚、他のワークステーションがそれぞれ公開した公開鍵の全てを公開鍵ファイル16dに格納しておく必要はない。例えばシステムに対して別に設けられた公開鍵ファイル・メモリに、各ワークステーションが公開した公開鍵を各ワークステーションに対応させてファイルしておく。そして情報通信が必要となったとき、その通信相手先の公開鍵を上記公開鍵ファイル・メモリから読み出して自己のワークステーションの公開鍵ファイル部18に格納するようにしても良い。

以上が暗号化処理部16の基本的な構成とその機能である。

次にイメージ照合部18について説明する。

このイメージ照合部18は、前記イメージ入力装置3から入力されたイメージ情報、例えば個人の顔のイメージを入力し、その個人同定を行なうものである。

第12図はこのイメージ照合部の概略構成を示すもので、18aはイメージ記憶部、18bは正規化

暗号化処理にあっては、上記信号系列M₁が通信原文であり、この通信原文から暗号化された通信文が信号系列N₁として求められる。また復号化処理にあっては上記信号系列M₁が暗号化通信文であり、この暗号化通信文から復号化された通信原文が信号系列N₁として求められる。

このような暗号化・復号化を担う鍵kが前述した公開鍵と秘密鍵であり、これらは対をなして設定される。

従ってワークステーションは、他のワークステーションから公開された公開鍵に従って通信情報をそれぞれ暗号化することはできるが、その暗号化された通信文を復号化し得るのは、その公開鍵と対をなす秘密鍵を知り得る特定のワークステーションだけとなる。

従って或る情報を暗号化して通信しようとするワークステーションは、通信相手先のワークステーションが公開した公開鍵に従って該通信原文を暗号化して通信する。そしてその通信情報は、秘密鍵を持つ通信相手先のワークステーションのみ

回路、18cは2値化(細線化)回路、18dは特徴データ抽出回路である。また18aはイメージデータを記憶したデータ記憶部であり、18fは検索回路、18gは照合回路、そして18hは出力部である。

イメージ記憶部18aは前記イメージ入力装置3を介して入力されたイメージ情報を記憶し、そのイメージ照合処理に供するものである。このイメージ記憶部18aに記憶されたイメージ情報に対して正規化回路18bは正規化処理し、また2値化回路18cは2値化処理する。具体的には、ここでは個人の顔のイメージからその個人同定を行なうべく、正規化回路18bはその顔の大きさを正規化している。この正規化された顔のイメージに対して2値化回路18cは、例えばエッジ線分検出、そのエッジ線分の細線化処理等を行なって該イメージの2値画像を求めている。

特徴データ抽出回路18dは、このようにして正規化・2値化されたイメージ情報をからその特徴データを抽出するものである。即ち、顔のイメージによる照合処理にあっては、例えば第13図に示

すように顔の輪郭を1つの特徴として抽出し、更にそのイメージ中の目、鼻、口等の特徴を抽出している。具体的には、顔の輪郭的特徴を分類されたコード情報として、また両眼間の距離と、口の大きさ、目と口との距離等を数値データとしてそのイメージの特徴として抽出している。

しかしてデータ記憶部18dには、予め各個人について求められた顔のイメージの特徴データが、例えば第14図に示すように登録されている。即ち、各個人毎にその個人名を識別名として上述した顔のイメージの特徴データが登録され、且つその顔のイメージ・データがポインタによって結ばれている。

検索回路18fは前記特徴データ抽出回路18dにて抽出された特徴データに基いて該データ記憶部18dを検索している。そしてその検索データは照合回路18gに与えられ、前記特徴データ抽出回路18dで求められた特徴データと照合処理されている。

この照合処理は、例えば特徴データ抽出回路

のバンドパスフィルタおよびA/D変換器等によって構成される。入力音声はこの音声入力回路19aにて、例えば30～3400Hzの周波数帯域の信号に制限され、12KHzのサンプリング周期で12ビットのデジタル信号に量子化される。

音響処理部19bは、例えば専用のハードウェアにより構成された积和回路からなる。そして基本的には前記音声入力回路19aと同期してバイブルイン的に高速動作する。

ここでの音響処理は、2種のバンドパスフィルタ群により実行される。その1つは16チャンネルのフィルタバンクで、このフィルタバンクを介して入力音声信号のスペクトルの変化が抽出される。今1つは、同じ帯域を4チャンネルに分割したグロスフィルタであり、このグロスフィルタを介して入力音声の音響的特徴が抽出される。

これらの2種類のフィルタ群(フィルタバンクとグロスフィルタ)は、例えば4次巡回形のデジタルフィルタとして構成される。そして、例えば10ms毎にそのフィルタリング出力を求めるも

18dで求められた入力イメージの特徴データを X_i (i は特徴の種別)、データ記憶部18eに登録されているイメージの特徴データを Y_i としたとき、

$$D = \sum_i |X_i - Y_i|$$

なる演算を行い、その演算結果Dの値が最も小さいものを、その個人として同定することによって行われる。この同定結果が出力部18hを介して出力される。

イメージ照合部18は、基本的にはこのようにして入力イメージを照合処理し、例えば該入力イメージの個人同定等を行なう。

次に音声認識部19について説明する。

音声認識部19は、例えば第15図に示すように構成される。音声入力回路19aは、前記音声入力部5から入力された音声信号、または公衆電話回線を介して前記通信装置12、13にて受信された音声信号を入力するもので、この入力音声信号を適当な信号レベルに增幅する増幅器や、帯域制限用

のとなっている。尚、この音響処理部の制御はマイクロプログラム方式にて行われる。

しかし前処理・認識部19cは、高速プロセッサ19d、パターンマッチング処理部19e、単語辞書メモリ19f、およびバッファメモリ19gによって構成される。

バッファメモリ19gは上記音響処理部19bにてフィルタリング処理された音声信号を入力し、例えば最大1.8秒分の音声データを蓄積するものとなっている。高速プロセッサ19dはこのバッファメモリ19gに格納されたデータに対して、音声区間検出、リサンプリング、ラベリング、遷移ネットワークによる認識処理、およびその総合論理判定処理の実行を行なっている。またこの高速プロセッサ19dにより、ホスト計算機との間の通信や該音声認識部19全体の動作制御が行われる。

この高速プロセッサ19dにて処理された音声データについて、パターンマッチング処理部19eは単語辞書メモリ19fに登録された単語音声の標準パターンデータとの間で複合類似度計算等のマッ

チング処理を実行し、その認識候補を求めている。

例えば認識対象となる音声単語は離散的に発声される。そこで高速プロセッサ19dは、例えば音響処理の際に10ms/sec毎に計算される入力音声エネルギーを用いて単語音声の入力区間を検出している。

具体的には第16図に示すように、背景雜音レベルと入力音声レベルとから適応的に計算される閾値 E_g を用い、入力音声信号レベルが上記閾値 E_g を一定時間以上継続して越えたとき、該閾値 E_g を越えた時点を音声単語の始端Sとして検出している。その後、上記入力音声信号のレベルが上記閾値 E_g を一定時間以上継続して下回ったとき、該閾値 E_g を下回った時点を音声単語の終端Eとして検出している。

ところで音声認識はパターン認識の一環として考え得る。しかし音声特有のパターン変動や、話者の性別、発声器官の形状、発声法等に起因する個人差、また話者自身が発生する雜音や周囲環境の雜音、更には電話音声の場合には公衆電話回線を経由したことによるレベル差や雜音の問題があ

る。このため、これらを考慮し、上述した変動要素を吸収して、如何に精度良く、安定に音声認識するかが問題となる。

そこでこの前処理・認識部19cではパターンマッチング法と構造解析法とを2段階に組合せ、ハイブリッド構造マッチング法と称される認識法を採用している。

即ち、上述したように単語音声区間が検出されると、先ずその音声区間(S, E)を15等分し、その16点をそれぞれリサンプル点とする。そして前述した如く音響処理された16チャネルの音声データ(スペクトル時系列)から上記各リサンプル点でのスペクトルを抽出する。尚、音声データのサンプル点と上記リサンプル点との間でずれがある場合には、リサンプル点の最近傍点のスペクトルを抽出すれば良い。

このリサンプル処理によって $16 \times 16 (= 256)$ 次元の音声パターン・ベクトルXを求める。即ち、第j(j=1, 2, 3, ..., 16)番目のリサンプル点を「j」とするととき、「j」での16チャネルのスペク

トルデータを

$$S(r_j) = (S_{1rj}, S_{2rj}, \dots, S_{16rj})$$

としてそれぞれ求め、これらの S_{irj} を並べ換えて

$$X = (S_{1r1}, S_{1r2}, \dots, S_{2r1}, \dots, S_{16r16})^t$$

なる音声パターンのベクトルXを求める。但し、tは行列の転置を示す。

このようにして求められた入力音声パターンベクトルXと、単語辞書メモリ19fに予め登録された単語音声の標準パターンとの類似度が、例えば複合類似度法によって計算される。

ここで単語辞書メモリ19fに予め登録された単語音声の標準パターンは、その単語カテゴリ ω_k について、

$$(\psi_{1k}, \psi_{2k}, \dots, \psi_{Lk})$$

$$(\lambda_{1k}, \lambda_{2k}, \dots, \lambda_{Lk})$$

但し、

$$(\lambda_{1k}, \lambda_{2k}, \dots, \lambda_{Lk})$$

として構成されている。尚、 ψ_{ik} , λ_{ik} はカテゴリ ω_k に属するパターンベクトルXの分散行列K

における固有ベクトルとその固有値である。このような単語辞書について、上述した複合類似度 $S(k)$ は

$$S(k) = \frac{\sum_{k=1}^L \lambda_{ik} \cdot (\mathbf{X}, \psi_{ik})^2}{\lambda_1 \cdot \|\mathbf{X}\|^2}$$

として計算される。尚、上式において $\|\mathbf{X}\|$ はベクトルXのノルムである。

このような複合類似度計算が全てのカテゴリについてそれぞれ行われ、上位に位置する類似度値と、それを得たカテゴリ名とが対にして求められる。

このような複合類似度法によるパターンマッチングによって、多くのパターン変動を救出した認識処理が可能となる。しかし類似パターンや雜音が加わったパターンでは、異なるカテゴリ間でその類似度値の差が小さくなることがある。

そこで前述したようにパターンマッチング法を補うものとして、以下の構造解析の手法を導入している。この構造解析は、単語音声を構成する音

の違いに着目して認識処理するもので、音素ラベル系列と音響的特徴系列の2つの時系列を利用していっている。

即ち、音素ラベル系列は、入力音声信号から10msac毎に計算される18チャンネルのスペクトルを用いて音素辞書との類似度を計算し、一定値以上の類似度を持つ音素のラベル付けして求める。尚、この音素ラベルは、例えば5つの母音と鼻音との6種類からなる。この際、音素辞書は、男声と女声に分けてそれぞれ準備しておく方が望ましい。

ここで比較的安定に発音される母音に比べ、子音を音素として個々にラベル付けすることが困難である。従ってその子音についてはその音響的な特徴をラベル付けし、これを特徴情報とする。具体的には、音響処理で求められる4チャンネルのグロスフィルタの出力と音声エネルギーとから音響的特徴を抽出する。このようにして特徴抽出されてラベル付けされる音響的特徴は、例えば第17図にグロスフィルタの出力の特徴と対比して示すように、無音性、無声性、摩擦性、破裂性、エネ

れを所定の閾値 θ と比較する。そして($S_1 < \theta$)の場合、これを無音としてリジェクトする。

また($S_1 \geq \theta$)の場合には、別の閾値 $\Delta\theta$ を用いて($S_1 - \Delta\theta$)以上の類似度を持つカテゴリを候補として抽出する。そしてその抽出されたカテゴリの数が1つである場合、これを認識結果として抽出する。また複数のカテゴリが抽出された場合には、前記遷移ネットワークによる解析結果を参照し、遷移ネットワークで受理されたカテゴリのみを抽出する。そしてその中で最大の類似度を持つカテゴリを認識結果として求める。

尚、閾値処理によって抽出されたカテゴリの中に、遷移ネットワークで受理されたものが含まれない場合には、判定不能とする。

以上のようにして複合類似度法によるパターン認識処理結果と、遷移ネットワークを用いた認識結果とを統合してその入力単語音声の認識が行われる。

第19図はこの音声認識部における単語音声の認識処理手続きの流れを示すもので、音声区間検

ルギ・ディップ等の12種類からなる。

しかして入力音声について求められた音素・音響ラベル系列は、前記音声区間(S, E)を含む範囲に亘って、各単語カテゴリ毎に作られた。例えば第18図に示す如き遷移ネットワークに入力される。

この遷移ネットワークの各ノード毎に、指定された音素ラベルや音響的特徴の有無をチェックする。そして無であればリジェクト、有であれば次のノードに遷移させ、その特徴系列が終了した時点で遷移ネットワークのゴールに到達した人力系列を受理し、そのカテゴリを求める。尚、系列のチェックの方向は、ネットワーク毎にその正逆を選択可能なものである。

総合判定論理は、前述した如くパターンマッチングによって順序付けられた候補カテゴリと、遷移ネットワークにより求められた認識結果とを統合して、その最終判定を行なうロジックである。

即ち、この総合判定論理は、パターンマッチングで求められた最大類似度を S_1 としたとき、こ

出処理の後、リサンブル処理してパターンマッチングを行い、同時にラベリング処理して遷移ネットワークによるチェックを行い、かかる後、これらの各認識結果を統合してその総合判定論理処理を行なうことが示される。このような処理が前記高速プロセッサ19dによる処理シーケンスの下で実行される。

ところで離散的に発声された単語音声ではなく、連続発声された音声中の単語を認識する場合には次のようにすれば良い。即ち、この場合には入力音声を種々の部分区間に分割し、その部分区間に単語識別を行なって単語類似度を求めるようにすれば良い。

具体的には、例えば第20図に示すように入力音声区間ににおける全ての分析フレーム間をそれぞれ部分区間の境界候補とし、該入力音声区間を複数の部分区間に分ける。この際、認識対象となる単語の総続時間長については最大時間長 D_{max} と最小時間長 D_{min} が設定できるので、その範囲内の部分区間だけを認識処理対象とすれば良い。

ここで第20図に示す例では、連続発声された音声の単語数が2個の場合を想定して2つの部分区間を求めている。しかし一般的には入力音声の単語数は不明であるから、2単語からN単語までが単語候補として存在すると仮定して部分区間をそれぞれ検出すれば良い。そして検出された各部分区間にについて単語類似度の計算を行い、その類似度結果の緊がり関係を相互に比較して最も信頼性の高い部分区間の境界を求めて、その境界によって区切られた部分区間の各単語認識結果を求めるようすれば良い。

然し乍ら、このようにして部分区間を求めて単語類似度計算を行なう場合、部分区間の数が膨大なものとなる為、処理の高適化が妨げられる。従って実際的には処理の高適化を考慮して、例えば入力単語数が2~5単語、1単語の継続時間長が128~840 msec、1回の発声における単語長の比が2.5以下、フレーム周期は16msec(8msec周期で2個に1個の単語を取出す)等の制限を加えて部分区間を検出するようすれば良い。

場合には84次元の継ベクトルとして与えられる。また母音パターンの場合には18次元の継ベクトルとして与えられる。

しかして特性値Kは、m個の学習パターンについて、その継ベクトル S_n と、この継ベクトル S_n を転置した横ベクトル S_n^T とを掛けさせて作成される行列の各成分を、上記四個の学習パターンに亘って平均化して求められる。従って特性値の要素数は上記ベクトルの要素数の2乗となる。

尚、このような処理によってそのカテゴリのパターン分布を反映した特性値Kを得るには、或る程度の量の学習パターンを必要とする。この為、学習パターン・メモリに予め所定数の学習パターンを蓄積しておくことが必要となる。

ところが母音の場合には18次元で最低6個のカテゴリの学習パターンを準備するだけで良いが、子音の場合には101カテゴリも存在し、しかも64次元のデータとして求める必要がある。この為、このままでは膨大なメモリ容量を必要とすることが否めない。

このようにすれば連続発声された音声中の単語をそれぞれ効果的に認識することが可能となる。

ところでこのような音声認識処理に供される辞書(単語辞書)の学習は次のようにして行われる。

この学習処理は、①母音パターンおよび子音パターンからその特性値を求める処理と、②その特性値に対する固有値と固有ベクトルを求める処理とに大別される。そしてこの固有値と固有ベクトルとを、その固有値の大きいものから順にN個求める。この処理は一般にKL展開と称されるものである。

先ず特性値を求める処理について説明すると、入力音声パターン(学習パターン)の特性値Kは、その学習パターンの継ベクトルを S_n としたとき、次のようにして求められる。

$$K = (1/H) \sum_{n=1}^H S_n S_n^T$$

ここに、

$$S_n = (S_{n1}, S_{n2}, \dots, S_{n84})^T$$

尚、この学習パターン S_n は、子音パターンの

そこで少數の学習パターンによってパターン分布を反映した特性値Kを得るべく、次のような特性値の更新処理を行い、逐次計算によってその特性値を次第にパターン分布を反映した形に改良していくようとする。

即ち、

$$K = K' + w S_n S_n^T$$

なる演算処理を繰返し実行するようとする。但し、wは特性値の更新時における重み係数である。この重み係数wは正負の値を取り、正ならば特性値行列の入力パターンに対する類似度を大きくし、逆に負ならば上記類似度を小さくする作用を有する。

また K' は S_n なる学習パターンを学習する前の特性値を示しており、Kは学習パターン S_n の学習によって更新された特性値を示している。

しかる後、このようにして求められた特性値に対して、その固有値と固有ベクトルを求める処理が行われ、この固有値と固有ベクトルとに基いて前述した複合類似度計算に用いられる標準バター

ンが作成される。

標準バターンは、上記特性値を K に展開することによって求められるものであり、例えばべき乗による K 展開によってその標準バターンが求められる。

今、特性値 K が固有値 $\lambda_1, \lambda_2, \dots, \lambda_n$ を持ち、これに対応する固有ベクトル e_1, e_2, \dots, e_n を持つものとする。この場合、その任意ベクトル u_0 は、上記固有ベクトル e_1, e_2, \dots, e_n の線形結合して

$$u_0 = \sum_{i=1}^n \alpha_i e_i$$

として表わされる。このとき、

$$K e_1 = \lambda_1 e_1$$

なる関係が成立することから、

$$\begin{aligned} K^s u_0 &= K^s \left(\sum_{i=1}^n \alpha_i e_i \right) \\ &= K^{s-1} \left(\sum_{i=1}^n \alpha_i \lambda_1 e_1 \right) \\ &= \dots = \sum_{i=1}^n \alpha_i \lambda_1^s e_1 \end{aligned}$$

$$v_{s+1} = K u_s$$

$$u_{s+1} = (v_{s+1}) / (b_{s+1}) \quad (s=0,1,2,\dots)$$

なる演算を実行するようとする。ここで (b_{s+1}) は、ベクトル (v_{s+1}) の絶対値が最大の要素である。このとき、

$$\begin{aligned} u_{s+1} &= (v_{s+1}) / (b_{s+1}) \\ &= (K u_s) / (b_{s+1}) \\ &\vdots \\ &= (K v_s) / (b_{s+1} \cdot b_s) \\ &\vdots \\ &= (K^{s+1} u_0) / (b_{s+1} \cdots b_s) \end{aligned}$$

となることから、これより $\lambda_1, b_{s+1}, e_1, u_{s+1}$ を求めることが可能となる。

このようにしてその絶対値が最大の固有値 λ_1 と固有ベクトル e_1 とを求めたら、次に同様にしてその絶対値が次に大きい固有値 λ_2 と固有ベクトル e_2 とを求める。

ここで

となる。

ここで

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$$

$$\begin{aligned} K^s u_0 &= \lambda_1^s [\alpha_1 e_1 + \sum \alpha_i \frac{\lambda_1^s}{\lambda_i} e_i] \\ |\lambda_1| &> 1 \quad (i=2,3,\dots,n) \end{aligned}$$

であるから、 S が十分大きくなると上式の第2項が0に収束することになる。

故に前述した式を

$$K^s u_0 = \alpha_1 \lambda_1^s e_1$$

と置換することができる。

このことは、 $(K^{s+1} u_0)$ と $(K^s u_0)$ の比が固有値 λ_1 であることを示している。また $(K^s u_0)$ は固有ベクトル e_1 に比例していることが示される。

ところでこのような理論に基く演算過程においては、その演算途中結果が直ぐにスケールアウトすることが多い。そこで u_0 を任意の、例えば単位ベクトルとし、

$$K' = K - \lambda_1 e_1 e_1^T$$

を考えると、

$$e_1^T e_1 = 0 \quad (i=2,3,\dots,n)$$

より、

$$\begin{aligned} K' e_1 &= K e_1 - \lambda_1 e_1 e_1^T e_1 \\ &= \lambda_1 e_1 - \lambda_1 e_1 = 0 \quad (i=1) \\ K' e_i &= K e_i - \lambda_1 e_1 e_1^T e_i \\ &= \lambda_1 e_i \quad (i \neq 1) \end{aligned}$$

となる。従って上記 K' は、

$$|\lambda_2| > \dots > |\lambda_r| > \dots > |\lambda_n| > 0$$

なる固有値を持つことがわかる。尚、ここでは e_1 は正規化されているとしている。

このような処理は、前記特性値を

$$K' = K - \lambda_1 e_1 e_1^T$$

として変換した K' に対して、上述した処理を繰り返し実行することによって達せられる。この処理によって絶対値の大きい固有値とそれに対応する固有ベクトルが順に求められ、辞書の学習が行われる。

第21回はこのような計算アルゴリズムに基いて

て実行される辞書の学習処理の手続きを示すものである。

次に文字認識部21について説明する。

この文字認識部21は、スキナ等によって読み取られた文字を認識する第1の文字認識ブロックと、タブレット等を介してオンライン入力される文字情報を認識する第2の文字認識ブロックとによって構成される。

この第1の文字認識ブロックは、例えば第22図に示すように、スキナ等によって読み取り入力された画像データを格納する画像メモリ21aと、この画像メモリ21aに格納された画像データ中から認識対象とする文字が記載された領域を検出する領域検出部21b、この領域検出結果に従って前記画像メモリ21aに格納された画像データ中から認識対象とする文字データを抽出する文字抽出部21c、そして標準パターン辞書21dに予め登録された認識対象文字の各標準文字パターンと、上記文字抽出部21cにて抽出された文字パターンとを個々に照合して文字認識する識別部21eとによって

の文字パターンの特徴を抽出し、その抽出した文字パターンと標準パターン辞書21dに登録された各文字の標準パターンとをパターンマッチングしている。そしてこのパターンマッチングによって照合の取れた標準パターンの文字カテゴリをその認識結果として求めている。

尚、パターンマッチングの手法は種々変形できることは云うまでもない。

ところでタブレット等を介してオンライン入力される文字情報を認識する第2の文字認識ブロックは、例えば第25図に示すように構成される。

この第2の文字認識ブロックは、タブレット等を介してオンライン入力される文字の筆記ストロークを示す位置座標の系列を順次検出する座標検出回路21hを備えている。

この座標検出回路21hにて検出された位置座標の時系列データは前処理回路21iに入力され、前記タブレット等における検出誤り等の微小な雜音が除去された後、座標系列記憶回路21jに順に記憶され、文字認識処理に供される。尚、この前処

て構成される。

この文字認識ブロックは、例えば第23図に示すようにFAX送信原稿用紙21f上の所定の位置に設定され、送信宛先が記入される文字枠21gに記載された文字を認識するものである。このような送信宛先が記載される原稿用紙21fは、送信原稿が複数枚からなる場合、その一番最初(1枚目)の原稿として用いられる。そしてこの1枚目の原稿の読み取り入力された画像データが文字認識処理の為に前記画像メモリ21aに蓄積される。

領域検出部21bは、予め定められているFAX送信原稿用紙21fのフォーマット情報から前記文字枠21gの位置情報を得、認識対象とする文字が記載される領域を検出するものである。文字抽出部21cはこの領域検出情報と、その画像情報の射影パターンの情報を用いて、例えば第24図に示すように前記文字枠21gに記載された文字の画像データを個々に抽出している。

識別部21eは、例えば特公昭49-12778号公報等に開示されるように、抽出された文字画像からそ

の文字パターンの特徴を抽出し、その抽出した文字パターンと標準パターン辞書21dに登録された各文字の標準パターンとをパターンマッチングしている。そしてこのパターンマッチングによって照合の取れた標準パターンの文字カテゴリをその認識結果として求めている。

また両数検出回路21kは、例えば筆記ストロークの途切れ(位置座標データの時系列の区切り)から、その文字パターンの筆記ストローク数、つまり両数を検出している。

しかしして認識処理部21eは、前記両数の情報を従って標準特徴パターンメモリ21aに登録された認識対象文字カテゴリの標準パターンの中から、該当する両数の標準パターンを選択的に抽出している。そしてこの標準パターンの各ストロークの特徴と座標系列記憶回路21jに記憶された入力文字パターンのストロークの特徴とを相互に比較(マッチング処理)している。審査決定回路21pはそのマッチング処理結果を判定し、入力文字パターンのストロークの特徴に該当するストロークを持つ認識対象文字カテゴリを、その認識結果として求めている。

つまりオンライン入力される文字パターンの並

記ストロークの特徴に従って、そのストロークの特徴を標準文字パターンのストロークの特徴とマッチング処理して上記入力文字パターンを認識するものとなっている。

尚、ストロークの特徴としては、筆記ストロークを折線近似したときの端点や交点、折点等の位置座標の情報を用いるようにすれば良い。

以上のような機能を備えた文字認識部21によって、スキャナ等を介して読み取り入力された文字情報や、タブレット等の位置座標入力装置を介してオンライン入力される文字情報をそれぞれ文字認識される。

次に图形認識部22について説明する。

この图形認識部22は、例えば第26図に示すように構成される。入力部22aは、例えば撮像入力された图形画像を記憶し、图形認識処理に供する。輪郭追跡部22bは、例えば線分の追跡方向を第27図に示すように8方向に分け、入力画像中の图形の輪郭を追跡したときにその追跡方向がどの向きであるかを順に求めている。具体的には、例

しきる後、セグメンテーション部22cにて輪郭の特徴点である端点、つまり曲りが急峻な点を検出し、その端点を中心としてその輪郭を分割する。そしてその分割された輪郭部分毎に辞書メモリ22cと照合し、その認識結果を求める。

以上の処理によって、第30図に示すように丸图形は端点が存在しないこと、三角图形は端点が3つ検出されること、四角图形は端点が4つ検出されること等から、これらの图形がそれぞれ識別認識される。この際、上記各端点がそれぞれ凸状であることや、端点を結ぶ輪郭が直線・曲線である等の情報を图形識別に利用しても良い。

これに対してイメージ認識部23は次のように構成される。

第31図はこのイメージ認識部23の概略構成を示すもので、原画像メモリ23a、2値化装置23b、処理画像メモリ23c、細線化装置23d、そしてコード変換装置23eによって構成される。

画像メモリ23aは与えられた認識対象イメージ画像を記憶するもので、2値化装置23bはこれを

えば第28図に示すように三角形の图形を右回りに追跡し、その追跡の向きの情報を、例えば「1,2,~2,3,4,~4,5,7,~7」なる方向コードの系列として求めている。

セグメンテーション部22cは、このようにして求められる方向コードの系列から、例えばその曲りの部分等の特徴点を抽出し、この特徴点に従って該图形の輪郭を複数の特徴部分に分割している。マッチング部22dはこのようにしてセグメンテーションされた图形輪郭の情報を、辞書メモリ22cに登録されている各種图形の特徴情報をマッチング処理して入力图形を認識するものとなっている。

例えば第29図に示す图形が与えられた場合には、その輪郭追跡によって求められる方向コードの系列から、例えば相互に隣接する3つの輪郭点(1-1),(1),(1+1)で方向コードの和を順に求め、これをその中央の輪郭点1における方向コードとして平滑化処理する。この平滑化処理によってノイズ成分の除去を行なう。

2値化処理して画像メモリ23cに格納している。この2値化レベルは、例えば2値化画像をディスプレイモニタしながら可変設定される。

しかし細線化装置23dは2値化されたイメージ画像を細線化処理してそのイメージを線图形化するものである。この細線化処理されたイメージ画像によって前記画像メモリ23cが書きえられて認識処理に供される。

コード変換装置23eは、例えば第32図に示すように構成され、先ずセグメント分割部23fにて上記細線化画像を複数のセグメントに分割している。このセグメントの分割は、例えば線图形をその端点や分岐点、交点にて分割することによって行われる。曲率変換部23gはこのようにして分割された複数のセグメントについて、それぞれその曲率を求めている。

直線・曲線分割部23h、曲線分割部23i、直折点分割部23j、および変曲点分割部23kは、上述した如く分割された各セグメントを、その曲率の情報を従って更に分割するもので、これらによって直折

点や直線と曲線との切替わり点、変曲点、曲線における半径変化点等がそれぞれ検出される。このようなセグメント分割と特徴点検出によって前記イメージ線図形を構成する各部の情報がそれぞれ抽出される。

近似情報作成部23aは、これらの分割されたセグメントおよびそのセグメント中の特徴点の情報を総合して前記イメージ図形を表現する情報、例えば各セグメントの始点および終点の位置座標、およびそのセグメントの種別を特定するコード情報を得る。

例えば入力イメージ画像が第33図(a)に示す如く与えられた場合、その入力画像中のイメージ線図形23bを直線化して抽出し、同図(b)に示すようにセグメント分割する。この例では、円図形と四角図形とが直線によって所謂半刺しにされたイメージ線図形23bが入力されている。しかしてこのイメージ線図形23bは、第33図(b)に示すようにその交点で分割され、2つの半円と2つのコの字状図形、および4つの直線にセグメ

ントとその位置座標として近似表現して認識するものとなっている。

さて音声照合部17は次のように構成されている。この音声照合部17は、音声入力した話者を個人認識(個人同定)するものであり、例えば第35図に示すように構成される。

即ち、音声入力部17aを介して与えられる音声は、音韻フィルタ17bおよび個人用フィルタ17cにてそれぞれフィルタリングされ、その音声特徴が抽出される。音韻フィルタ17bの複数のチャンネルの各帯域は、例えば第36図(a)に示すように音声周波数帯域を等分割して設定されている。このようなフィルタ特性を備えた音韻フィルタ17bによって入力音声の音韻特徴を示す特徴パラメータが抽出される。尚、各チャンネルの帯域幅を、音声周波数帯域を対数閏数的に分割設定したものとしても良い。

これに対して個人用フィルタ17cの複数のチャンネルの各帯域幅は、第36図(b)に示すように音声周波数帯域を指数閏数的に分割して設定さ

ント化される。

曲率変換部23gは、第34図に示すようにセグメント分割された各セグメントの曲率を求めており、前記直線・曲線分割部23h、曲線分割部23i、屈折点分割部23j、および変曲点分割部23kはその曲率変化点から各セグメントの特徴点を検出している。具体的には第34図(a)に示す例では2つの直線の屈折点における曲率が急峻に増大することから、その曲率の変化から屈折点を検出することが可能となる。また第34図(b)に示す例では直線から曲線への変化部分で曲率の変化が検出されるので、この曲率の変化からその特徴点を検出することができる。

同様にして第34図(c)、(d)に示す例でも、その曲率の変化点から、そのセグメントにおける特徴点を検出することが可能となる。

このようにしてイメージ認識部23では、与えられたイメージ図形をセグメント化し、各セグメントの特徴点を検出している。そして接イメージ線図形を複数のセグメントの各種別を示すコード情

報とその位置座標として近似表現して認識するものとなっている。このようなフィルタ特性を備えた個人用フィルタ17cによって、前記入力音声の低域から中域にかけての音声特徴が、高域側の特徴に比較して多く抽出されるようになっている。そしてこれらの各チャンネルのフィルタ出力が個人用の特徴パラメータとして求められている。

しかして単語認識部17dは、前記音韻フィルタ17bを介して求められた音韻特徴パラメータから、その入力音声が示す単語を単語辞書17eを基盤として認識するものである。この単語認識の機能は前述した音声認識部19と同様であり、該音声認識部19の機能をそのまま利用するようにしても良い。そしてこの単語認識結果に従って個人辞書17fの個人照合に供される辞書が選択される。この個人辞書17fは、話者照合の対象とする個人が予め充填した特定の単語の前記個人用フィルタ17cによる分析結果を、その単語毎に分類して登録したものである。

しかし話者照合部17gは、個人辞書17fから選択された該当単語の各特徴パラメータと、前記

個人用フィルタ17cにて求められた入力音声の特徴パラメータとの類似度を計算し、その類似度値を所定の閾値でそれぞれ弁別している。そしてそれらの弁別結果を相互に比較して、例えば類似度値が最も高く、次に高い類似度値との差が十分にある特徴パラメータを得た個人カテゴリを該入力音声の危険者であるとして個人同定している。

ここで個人用フィルタ17cの特性について更に詳しく説明すると、前述したように音韻特徴フィルタ17bとは異なる特性に設定されている。この音声の個人性の識別性について考察してみると、その識別性は、例えば

$$F\text{比} = (\text{個人間分散}) / (\text{個人内分散})$$

として与えられるF比によって評価することができる。

今、音韻フィルタ17bに設定されたフィルタ特性の各チャンネル出力のF比について検討すると、第37図に実線で示す指數関数的な傾向を示す。これ故、従来では専ら高域側の音声特徴情報を利用して個人照合を行なっている。

められる。この結果、高域側の音声特徴のみならず、中域における音声特徴をも積極的に利用して個人照合を行なうことが可能となり、その照合精度の向上を図ることが可能となる。

即ち、この音声照合部17では、入力音声の単語認識に供する特徴とは別に、フィルタバンクの工夫によりその個人性が顕著に現われる特徴情報を抽出している。この結果、入力音声に対する音韻認識とは独立にその話者に対する個人同定、つまり個人照合を高精度に行なうものとなっている。

次に音声合成部28について説明する。

音声合成部28は、第38図に示すように判別器28a、復号器28b、規則パラメータ生成装置28c、および音声合成器28dを編んで構成される。

判別器28aは入力されたコード列が文字列であるか、或いは音声合成の為の分析パラメータを示す符号列かを判定するものである。この情報判別は、例えば入力コード列の一一番最初に付加された識別情報を判定することによって行われる。そして分析パラメータであると判定した場合には、そ

しかし音声の高域側の特徴だけを用いるよりも、全周波数帯域の音声特徴を用いて個人同定が可能であれば、その照合精度が更に向上すると考えられる。即ち、全周波数帯域においてF比の値が1以上となり、個人間分散が個人内分散を上回れば、更に精度の高い個人照合が可能となる。

そこでここでは、前述したように個人用フィルタ17cの特性を指數関数的に定め、個人性の特徴が顕著である高域側については大雑把に特徴抽出し、低域側のチャンネル調当て数を増やすことによって該低域側の音声特徴を細かく抽出するようしている。

具体的には各チャンネルのF比の変化が指數関数的な傾向を示すことから、低域側チャンネルの帯域幅に比較して高域側チャンネルの帯域幅を指數関数的に増大させたフィルタバンクを構成し、これを個人用フィルタ17cとしている。

このように構成されたフィルタ17cの各チャンネル出力によれば、そのF比は第37図に破線で示すようになり、中域でのF比の大幅な向上が認

の符号列を復号器28bに与え、これを復号処理してその音韻パラメータと韻律パラメータとをそれぞれ求めている。

また文字列と判定した場合には、その文字列データを規則合成パラメータ生成装置28cに与え、その音韻パラメータと韻律パラメータとの生成に供している。

音声合成器28dは、このようにして復号器28bまたは規則合成パラメータ生成装置28cにて求められた音韻パラメータと韻律パラメータとに従い、音波を声道近似フィルタを介して処理して合成音声波を生成している。

ここで規則合成パラメータ生成装置28cについて更に説明すると、該装置28cは第39図に示す如く構成されている。文字列解析部28eは言語辞書28を参照して入力文字列中の単語を個々に同定し、その単語についてのアクセント情報や単語・文節境界・品詞・活用等の文法情報を求めている。そしてこの解析結果に対して音韻規則、および韻律規則がそれぞれ適用され、その韻律情報を生成

される。

ここで音韻規則は、解析された単語の読みの情報を与えると共に、単語の連接によって生じる連濁や無声化等の現象を実現し、その音韻記号列を生成するものである。音声パラメータ生成部26gはこの音韻記号列を入力し、その音節単位に従ってCVファイル26hから音節パラメータを順次求めて補間結合している。この音声パラメータ生成部26gにて上記音韻記号列から音韻パラメータ系列が生成される。

また韻律規則は、単語・文節境界等の文法情報に従って発話の境界や息継ぎ位置を決定し、各音の離続時間長やポーズ長等を決定するものである。同時にこの韻律規則により、各単語の基本アクセントをベースとし、文節アクセントを考慮した韻律記号列が生成される。韻律パラメータ生成部281はこの韻律記号列を入力し、ピッチの時間変化パターンを表わす韻律パラメータ列を生成している。

一方、入力コード列が音声合成のための分析パラ

メータのケプストラム係数C₀に応じてレベル制御されて対数振幅近似ディジタルフィルタ281に入力される。この対数振幅近似ディジタルフィルタ281は前記音韻パラメータのケプストラム係数C₁、～C_nに従って声道特性を近似する共振回路を構成し、上記音韻波をフィルタリング処理するものである。この対数振幅近似ディジタルフィルタ281にて前記音韻パラメータおよび韻律パラメータで示される音声データが合成出力される。

そして対数振幅近似ディジタルフィルタ281にて合成された信号は、D/A変換器28uを介した後、LPF28vを介してフィルタリングされて合成音声信号（アナログ信号）として出力される。

以上のように構成された音声合成部28にて、入力データ系列からそのデータ系列が示す音声が規則合成されて出力される。

次にイメージ合成部27について説明する。

イメージ合成部27は、第42図に示すように制御演算部27a、ディスプレイファイルメモリ27b、イ

メータを示す符号列である場合、前記復号器26bは次のように機能している。

即ち、分析パラメータの符号列がCVファイルのケプストラム係数を示す場合、その符号列26hは一般に第40図に示すようにパラメータP（ピッチ）とC₀、C₁、～C_n（ケプストラム係数）に対してピット割当てがなされて情報圧縮されている。そこで復号器26bではパラメータ変換テーブル28nを用い、上記情報圧縮された分析パラメータを音声合成器28dに合せたピット数に変換・復号している。例えば各パラメータをそれぞれ8ビットに変換し、音韻パラメータ列（ケプストラム係数）とその韻律パラメータ列（ピッチ）とをそれぞれ求めている。

音声合成器28dは、例えば第41図に示すように有声音源28qと無声音源（M系列発生器）28rとを備え、入力される韻律パラメータ列のピッチデータPに従って有声音波（P≠0）、または無声音波（P=0）を選択的に発生している。この音波は前記増幅器28sに入力され、前記音

メージ合成回路27c、イメージメモリ27d、そして必要に応じてディスプレイ27eを備えて構成される。尚、このディスプレイ27eは、該ワークステーションについて準備された前記ディスプレイ部10であっても良い。

イメージ合成回路27は、専用の制御演算部27aの制御の下でディスプレイファイル27bに蓄積されているベクトルや多角形・円弧のパラメータを取出し、それによって示される線図形を発生してイメージ・メモリ27dの指定されたアドレスに蓄積している。このイメージ合成回路27のイメージ発生機能によってイメージメモリ27d上に指定された線図形イメージが構築される。そしてこの線図形イメージは、制御演算部27aの制御の下で前記ディスプレイ27eにて表示されてモニタされる。

またイメージ発生回路27bは、イメージ発生に対する特殊処理機能と塗潰し処理機能とを備えている。この特殊処理機能は、例えば複数のイメージ図形の重なりに対して図像の消去を行なったり、クリッピング処理を行なう等の機能からなる。ま

た塗潰し機能は、イメージ图形の部分領域を指定された色を用いて塗潰す処理からなる。

このようなイメージ合成回路27bの機能によって、種々のイメージ图形が作成され、またその合成処理等が行われる。

ところで上述した如く発生したイメージ图形と自然画との合成は次の2つに大別される。その1つは、例えば風景写真等の自然画を背景として、その中に演算処理によって求められたイメージ画像を埋め込み合成する処理であり、他の1つのは制御演算部27aが内部モデルとして持っている或る平面イメージ内に自然画を埋め込み合成する処理からなる。

ここで前者の自然画中にイメージ画像を埋め込み処理する場合には、例えば第43図にその概念を例示するように、制御演算部27aが発生する图形中に「透明色」を示すコードを与えておき、これを自然画に対して重ね合せて合成することによって達せられる。すると「透明色」コードが与えられた西像領域は、自然画の情報がそのまま表示

されることになり、その他の部分は制御演算部27aが発生した图形が表示されることになる。この結果、自然画を背景としたイメージ合成が実現されることになる。この手法はオーバーレイと称される。

これに対して第44図にその概念を示すように画像メモリ内に自然画を荷込んでおき、その上(手前)に制御演算部27aが発生した图形を荷込んで行くようにしても良い。この手法はエバッファ法と称されるものであり、前述したオーバーレイ法と共に比較的簡単に実現することができる。

ところで制御演算部27aの内部モデルとして示される平面内に自然画を嵌め込み合成する後者の場合には、次のようにして高速処理される。

平面上にある自然画を、3次元空間内の任意の方向を向いている平面に埋め込む為に必要な座標変換は次式で与えられる。

$$u = \frac{C_1 X + C_2 Y + C_3}{C_4 X + C_5 Y + C_6}$$

$$v = \frac{C_7 X + C_8 Y + C_9}{C_4 X + C_5 Y + C_6}$$

但し、X、Yは表示面での座標であり、u、vは自然画での座標である。

この座標変換処理をそのまま実行しようとすると、1画素を表示する毎に6回の乗算と2回の除算が必要となり、膨大な計算量と計算処理時間を必要とする。

そこでここでは、次のような中間座標(s, t)を介して上述した演算を2回の変換処理に分解して実行するものとなっている。この演算処理は、例えばアフィン変換を利用して高速に実行される。

$$u = (\alpha_1 s + \alpha_2 t + \alpha_3) / t \quad (1)$$

$$v = (\alpha_7 s + \alpha_8 t + \alpha_9) / t \quad (2)$$

$$s = C_5 X - C_4 Y \quad (3)$$

$$t = C_4 X + C_5 Y + C_6 \quad (4)$$

即ち、上述した第(1)式を用いて透視変換を行い、その後、第(2)式を用いて2次元アフィン変換を行なって任意の平面への透視変換を高速に行なうも

のとなっている。

ここで、第(1)式の分母は座標tそのものであるから、従来より知られているアフィン変換回路を若干改良するだけでその演算を高速に実行することができる。

このようにしてイメージ合成部27では種々のイメージ合成処理を高速に実行するものとなっている。

次にデータベース部32について説明する。

データベース部32はコードやイメージ、音声等の各種のデータを整理して格納し、これを種々の応用システムに供するものである。第45図はこのデータベース部32の概略構成を示すもので、コマンドの解析処理等を実行するインターフェース部32a、データベースの検索処理等を実行するデータ操作部32b、種々のデータを格納する記憶媒体としての磁気ディスク装置32cや光ディスク装置32d、そしてその附加機能部32eとによって構成される。

種々のデータは、そのデータの種別に従って投

数のリレーションに分類整理され、各リレーション毎にそれぞれ登録されてデータベースが構築されている。

以下、このデータベース部32を、その論理構造、蓄えられるデータ、物理構造、および付加機能の4つに分けて説明する。

論理構造とはこのデータベース部32を応用システム側から見た場合、種々のデータがどのように蓄積されているかを示すものである。ここではリレーションナル・モデルに従った論理構造として、例えば第46図に示すような表のイメージとしてデータが取扱われるようになっている。

表(リレーション)には幾つかの欄(アトリビュート)が設けられており、これらの各欄に所定の単位のデータがそれぞれ格納される。データの単位(タップル)は、各欄に格納すべき1組の値として定められる。このようなタップルを格納した任意個数のアトリビュートによって1つのリレーションが構築される。

しかし、このモデルにあっては、リレーション

名を指定し、その各アトリビュートの値をそれぞれ与えることによってデータベースへのデータの格納が行われる。またデータベースの検索は、リレーションおよびアトリビュートを指定し、そこに格納されている値が指定された値、または別のアトリビュートに格納されている値との間で所定の条件を満すか否かを判定し、その条件を満すタップルを抽出することによって行われる。

この検索条件は、それらの値が等しい、等しくない、小さい、大きい等として与えられる。この際、複数のアトリビュートについてそれぞれ検索条件を指定し、その条件判定結果を論理処理(アンドやオア等)して行なうことも可能である。更には、複数のリレーションを指定し、或るリレーションの或るアトリビュートの値が他のリレーションの或るアトリビュートの値に等しい等の条件により、複数のリレーション中から所定のタップルを求めるようなデータベース検索も可能である。

またデータベースからのデータ削除は、基本的には上記検索と同様に行われるが、タップルを抽

出することに代えて、そのタップルを抹消することによって行われる。

更にデータ更新も同様であり、得られたタップルの指定されたアトリビュートの値を変更し、これを格納することによって行われる。

また各リレーションには、各アトリビュート毎にデータの読み出し、追加、変更が許可された人の情報(人名や担当者コード)等が記入され、データ保護の対策が講じられている。尚、このデータ保護対策をアトリビュート毎に行なうことに代えて、リレーション単位で行なうことも可能である。尚、ここに記載される人の情報は複数であっても良い。

しかし第46図に示すリレーションの例では、文字列としてそのデータが示されているが、各リレーションに蓄積されるデータは単なるビット列であっても良い。つまりリレーションに蓄積されるデータは文字列は勿論のこと、イメージ情報や音声情報等であっても良い。

さてこのデータベースに蓄積されるデータは、

上述した第46図に示す「個人スケジュール」のリレーションを始めとして、例えば第47図に示すような「住所録」「個人の仕事とその代行者」「操作履歴」「人事」「会議室」「会議室予約」「会議」等の種々のリレーションからなる。

この例に示されるようにリレーションは、主に個人用として用いられるものと、多くの利用者によって共通に利用されるものとからなる。そして個人用のリレーションは各個人が使用するワークステーション毎に設けられ、また共通リレーションは複数の利用者にとって共通なワークステーションに設けられる。

尚、共通のワークステーションとは必ずしもそのハードウェアが他のワークステーションと異なることを意味しない。また個人用のワークステーションが共通のワークステーションを兼ねても良いことも勿論のことである。更には共通のワークステーションは1台に限られるものではなく、システムの階層レベルに応じて複数台設けられるものであっても良い。要するに、複数のワークステ

ーションから容易に特定することのできるものとして普通のワークステーションが設定される。

ここで第46図に示した「個人スケジュール」リレーションのデータ構造について簡単に説明する。

このリレーションからは、そのリレーション名が「個人スケジュール」であり、「△△△△」によって作成されたことが示される。このリレーション作成者「△△△△」は該リレーションに対して全てのデータ操作が許可される。

またこのリレーションに付加されたデータ保護機能によれば、データの読み出しは全員に許可されており、データの追加は「〇〇〇〇」と「技術部に所属する者」に対してのみ許可されている。尚、この「技術部に所属する者」は、例えば「人事」のリレーションを参照する等して求められる。またデータの変更は「人レベル」の値が「5」以上のものに対してのみ許可されている。この「人レベル」とは人事リレーションに関するものであり、例えば（部長：8）（次長：7）（課長：6）

第2ページに使用中のページ情報をそれぞれ格納する。

このリレーションの一覧表によって、データベース中における種々のリレーションの所在が示される。

例えば第9ページおよび第11ページに格納された実データは、第5ページに格納されたリレーションのアトリビュート（主アトリビュート）に基づき、第10ページに格納されたインデックスページの情報を従ってソートされるようになっている。このインデックスページの情報は、アトリビュートの値が幾つから幾つ迄のものがどのページに格納されているかを示すものである。

この主アトリビュート以外のアトリビュートによりデータ検索する場合には、そのアトリビュートについて第20ページのサブ・インデックスを絞り出し、先ず第21ページや第22ページに示されるサブデータを得る。このサブデータにはアトリビュートの値と前述した主アトリビュートの値のみが入っており、ここで求められるアトリビ

ュートの値を用いて実際のデータが求められる。

更にこのリレーションには、「開始時刻」「終了時刻」「種類」「名称」「場所」等のアトリビュートが設定され、そのそれにデータが書き込まれるようになっている。

次にこのデータベース部32における上述した各種のリレーションを実際記憶するための物理構造について説明する。

情報蓄積部（記憶部）は大量データを蓄積し、その任意の部分を比較的高速に読み書きすることができ、価格的にきぼど高価でないものとして前述した磁気ディスク装置32cや光ディスク装置32gが用いられる。

この情報蓄積部へのデータベースの蓄積は、該情報蓄積部の記憶領域を特定の大きさ（例えば数キロバイト程度で、タップル長や計算機の速度等に応じて定められる）毎に区切り、各々をページとして管理して行われる。そして第48図に示すように、例えば第0ページにページ管理の情報を、第1ページにリレーション一覧表の情報を、また

ユートの値を用いて実際のデータが求められる。

尚、例えば画像データや音声データのようにその実データの量が膨大であり、その中の幾つかのピット誤りが問題とならない場合には、これらの実データを光ディスク装置32d等の別の安価な情報記憶装置にファイルするようにしても良い。この場合には、第9ページや第11ページ等の実データ用ページには、その旨とその装置での実データの格納位置情報を記憶しておくようにすれば良い。

しかしてこのように構築されたデータベースに対する付加機能は、例えば不要データの自動廃棄等からなる。この不要データの自動廃棄は、リレーションの付加情報として【廃棄：可／不可】【廃棄の方法】等を与えておき、所定の間隔でリレーション毎の消去コマンドを動作させて行われる。

尚、タップルの消去は、例えば会議情報についてはその終了時刻が現在の時刻より前であるか否か等を判定して行なうことが可能である。従って

このようなタップルの消去については、格別の機能追加は不要である。

また付加機能の他の重要な機能としてデータの保全がある。このデータの保全機能は、例えばハードウェアの故障や停電等に原因してデータが不正（でたらめになったり失われたりすること）となることを防ぐものである。具体的にはこのデータの保全機能は、情報の二重化や磁気テープへの書き出し等によって実現される。

このようにデータベース部32では、種々のデータをリレーション毎に分類整理し、且つページ単位に管理して種々の応用システムに供するものとなっている。

次に作業環境データ収集部25について説明する。この作業環境データ収集部25は、該ワークステーションに対する過去の操作履歴のデータを収集し、これに基く操作ガイドを行なうものである。

ここで作業環境データ収集部25には、例えば第49図に示すように当該情報処理システムが持つ機能に対応するコマンドと、他の情報システムが

持つコマンドがシステムAのものであるか否かが調べられる。そして入力コマンドがシステムAのコマンドであると解析されると、コマンド解析部25bは該入力コマンドをコマンド実行部25dに与え、そのコマンドに基く所定の動作を実行させている。

一方、入力コマンドがシステムAのものでない場合には、他のシステムのコマンドに該当するか否かが調べられ、対応付けされているコマンドが存在する場合には、その対応コマンドを画面表示部25eにて表示する。つまり他のシステム（システムB）で用いられているコマンド、例えば“DEL”である場合には、これに対応するシステムAのコマンド“DELETE”を求め、これを操作ガイドとして画面表示部25eに表示することになる。

尚、入力コマンドに該当するコマンドがコマンド対応テーブル25cに存在しなかった場合には、画面表示部25eにてコマンドエラーメッセージの表示を行なう。

持つ機能に対応するコマンドとを対応付けるコマンド対応テーブルが設けられている。

具体的には当該情報処理システムをA、他の情報処理システムをB、C、D、…としたとき、システムAにおけるコマンド“DELETE”に対応する他のシステムのコマンドが“DEL”“ERASE”“REMOVE”であることが、該コマンド対応テーブルによって示されるようになっている。

第50図は利用者により入力されたコマンドを解析し、所定の動作および各種ガイダンスを実行する作業環境データ収集部25の概略構成を示すものである。

この作業環境データ収集部25では、先ずコマンド入力部25aから入力されたコマンドをコマンド解析部25bに与え、コマンド対応テーブル25cを参照して解析している。具体的には第51図に示す手順の流れに従って入力コマンドがコマンド対応テーブル25cに登録されているかを調べている。即ち、コマンドが入力されると、先ずその人

具体的には次のようにしてそのコマンド入力に対する処理が行われる。今、システムB、Cの操作経験の利用者が初めてシステムA（当該情報処理システム）を操作するものとする。ここで利用者がコマンドを入力してデータ“ABC”を消去する場合、従来ではシステムAの取扱い説明書に従ってデータ消去の為の“DELETE”なるコマンドを探し、これを入力することが必要となる。

しかしここでは、その利用者は過去の経験に従って、例えばシステムCで用いていたデータ消去コマンド“ERASE ABC”を第52図(a)に示すように入力する。

すると作業環境データ収集部25ではこの入力コマンドを解析し、明記コマンド対応テーブル25cから入力コマンド“ERASE”に対応するシステムAのコマンド“DELETE”を求め、これをガイドとして表示することになる。この結果、利用者はシステムAを初めて操作する場合であっても、そのデータ消去のコマンドが“DELETE”であることを知り、そのコマン

ドをガイドに従って入力することにより、そのデータ消去を行なうことが可能となる。

またファイル名のリストを表示するべく、第52図(b)に示すようにシステムBにおけるコマンド“DIR”を入力した場合には、同様にして該システムAにおける対応コマンド“CATA”が求められ、ガイド表示される。この結果、このガイドに従ってコマンド“CATA”を入力することによって、そのファイル名のリストが表示される。

このようにこの作業環境データ収集部25の機能を活用することにより、過去の操作経験のあるシステムで用いられていたコマンドの入力によって、そのシステムにおける対応コマンドがガイド表示される。従ってシステム利用者は、過去に得た知識を最大限に利用してシステムを操作することが可能となる。そして当該情報処理システムのコマンドを容易に知ることが可能となる。従ってその都度、当該情報処理システムの操作説明書を調べる等の煩わしさから解放される。故に、システム

れて、その処理が実行されることになる。従って、更にその操作性の向上を図ることが可能となる。

尚、対応コマンドはシステムの種類に応じて何種類存在しても良いものである。要はコマンド対応テーブル25cに対応付けてそれぞれ格納しておけば良い。またコマンドは上述した文字列形式に限定されないことも云うまでもない。

次にこの作業環境データ収集部25におけるシステム習熟度のデータ収集について説明する。この作業環境データ収集部25の内部に、このシステム習熟度のデータ収集処理を行なう為のハードウェアとして外部記憶装置と翻譯装置がおかれる。

第55図はシステム習熟度のデータ収集処理を示す流れ図である。

利用者がその識別コード(ユーザ番号やパスワード等)を入力すると、作業環境データ収集部25はその識別コードに対応する習熟度表を外部記憶装置から求め、装置内部にセットする。この習熟度表は各利用者がシステムの様々な利用機能に対してどの程度習熟しているかを格納したもので、

その操作の習得に要する時間を大幅に短縮することができる等の効果が期待できる。

尚、入力コマンドに対応するコマンドを決め、これをガイド表示したとき、その合否の判定入力を受けて、そのコマンドを実行するようにしても良い。

即ち、第51図にその手続きの流れを示し、第54図にその表示例を示すように他のシステムの消去コマンド“ERASE”し、これに対応するシステムAの消去コマンド“DELETE”が求められたとき、これが正しいか否かを問合せる。そして正(Y)なる指示入力があったとき、その入力コマンドが“DELETE”を示していると判定し、これをコマンド実行部25dに送ってその処理を実行させるようとする。

このようにすれば、コマンドの対応関係がガイド表示されると同時に、その入力コマンドに従って所要とする処理が実行されるので、改めて正しいコマンドを入力し直す必要がなくなる。つまり入力コマンドの対応コマンドへの自動変換が行わ

れば第56図に示す如く構成されている。

即ち、この習熟度表は各利用機能に対してその利用頻度、最終利用年月日時、ユーザが申告した該機能に対する習熟クラス、該機能を前回利用した際の習熟度クラス、更には該機能の複雑度の情報等によって構成されている。

ここで複雑度とは該当利用機能が専門知識を要求する程度となり、また基本機能より高級機能になる程高くなるものである。

しかしてこのような習熟度表は各利用者毎に設けられ、外部記憶装置にそれぞれ記憶されている。尚、システムを初めて利用する利用者に対しては、識別コードの新規設定によりその利用者に対する習熟度表が作成され、外部記憶装置に登録される。

尚、外部記憶装置には、例えば第57図に示すように上述した習熟度表に加えて、前記習熟度クラスに対応した利用機能毎のメッセージが登録されている。このメッセージは習熟度のクラスが低い程、その背景説明を含む判り易い説明となっている。また習熟度の高いクラスほど、簡略な説明

と専門的な機能の紹介を含んだ高度な内容となっている。

また習熟度のクラスは、例えば
A: 初級者クラス
B: 中級者クラス
C: 賢熟者クラス

のように分類設定される。

しかして入力された識別コードに対応した習熟度表が求められると、次にその利用機能を利用者に選択させる為のメニューが表示される。このメニューに対して利用者は、例えばその利用機能に対応する番号等を入力する。すると制御装置ではその入力情報が終了信号か利用機能の選択信号かを判断し、利用機能選択信号の場合には次のように動作する。

即ち、利用機能選択信号が入力されると、先ずその利用者に関する前記習熟度表を参照し、選択された利用機能に対する利用頻度や最終利用年月日時、中級者度クラス等の情報が求められる。そしてこれらの情報に従って重み付け処理を施し、

される。

ここで $G_1 [X_1]$ は、 $X_1 = A$ のとき Y_1 なる値を取り、 $X_2 = B$ のとき Y_2 なる値を取ることを意味する。また $(T_c - T_0)$ は、最終利用年月日時から現在までの日数を時間換算したものである。

しかしてクラス判定は、上述した判別関数 F_r の値により次のようにして行われる。

$$\begin{aligned} F_r < N_1 & \cdots A \text{ クラス} \\ N_1 \leq F_r < N_2 & \cdots B \text{ クラス} \\ N_2 \leq F_r & \cdots C \text{ クラス} \end{aligned}$$

尚、判定閾値 N_1 、 N_2 は実験等に基いて適切に定められる。

このようにして習熟度クラスが決定されると、その決定された習熟度クラスに対応し、且つ前述した如く指定された利用機能に該当するガイドメッセージやエラーメッセージを外部記憶装置から求める。

しかる後、今回決定された習熟度クラスと、前記習熟度表に格納されている前回の習熟度クラス

現在の習熟度クラスの決定が行われる。

この習熟度クラスの判定は、例えば利用頻度を P_1 、最終利用年月日時を T_0 、現在の利用年月日時を T_c 、利用者中告習熟度クラスを X_1 、前回利用習熟度クラスを X_2 (A 、 B 、 C)、複数度を P_c 、そして判別関数を F_r としたとき、

$$\begin{aligned} F_r = K_1 P_1 + K_2 (T_c - T_0) \\ + K_3 G_1 [X_1] \\ + K_4 G_2 [X_2] + K_5 P_c \end{aligned}$$

として求められ。但し、上式において K_1 、 K_2 、 K_3 、 K_4 は、実験等によって適切な値に設定される定数である。また上記 G_1 、 G_2 は

$$G_1 = \begin{Bmatrix} A & B & C \\ Y_1 & Y_2 & Y_3 \end{Bmatrix}, G_2 = \begin{Bmatrix} A & B & C \\ Z_1 & Z_2 & Z_3 \end{Bmatrix}$$

であり、 Y_1 、 Y_2 、 Y_3 、 Z_1 、 Z_2 、 Z_3 は、 A 、 B 、 C に対する評価並みである。これらの評価並みは

$$Y_1 < Y_2 < Y_3, \quad Z_1 < Z_2 < Z_3$$

なる関係を有し、実験等によって適切な値に設定

とを比較照合する。そして習熟度クラスに変更がある場合には、その習熟度に変更がある旨を示すメッセージを前記ガイドメッセージ等に付加して表示する。

この習熟度クラス変更のメッセージは、例えば第58図に示すような4種類のメッセージからなる。そしてそのクラス変更の形態に応じて求められ、前記ガイドメッセージ等と共に表示される。利用者はこのようにして表示される各種メッセージに従ってその処理操作を行なうことになる。

具体的には作成データをファイルに格納する利用機能に対して、その利用者が初級者クラス (A クラス) と判定されると第59図に示す如きメッセージが表示される。そしてこのメッセージにも拘らず利用者が情報入力を誤った場合には、例えば第60図に示すようなエラーメッセージの表示が行われ、その利用機能に対する操作のガイドが行われる。

またその利用者の習熟度が中級者クラス (B クラス) と判定された場合には、第61図に示す如

きメッセージが表示される。そしてこのメッセージにも拘らず利用者が情報入力を誤った場合には、例えば第62図に示すようなエラーメッセージの表示が行われ、その利用機能に対する操作のガイドが行われる。同様にその利用者の習熟度が習熟者クラス(Cクラス)と判定された場合には、第63図に示す如きメッセージが表示され、情報入力の誤りがある場合には、例えば第64図に示すようなエラーメッセージの表示が行われてその利用機能に対する操作のガイドが行われる。

しかして上述した如く表示したガイドメッセージの空欄に対してデータ入力が行われると、制御装置は前述した如く求めている該当利用者の習熟度の該当利用頻度を(+1)すると共に、最終利用年月日時および前回利用習熟クラスの更新を行なう。そして該利用機能の実行を促すと共に、該利用機能が終了したものと看做して前述した利用機能選択の為のメニュー表示動作に戻る。

ここで再び利用機能選択信号が入力されると、上述した処理を再び繰返して実行することになる。

認識・イメージ合成、更には音声やイメージ等の圧縮・伸長処理等の種々のメディア変換機能を備えている。

このようなメディア変換を情報通信する通信相手端末の種別に応じて選択制御し、通信相手端末が処理可能なメディアに変換して通信出力する機能が本発明の特徴とするところであり、その処理は主として出力形態選択部24にて行われる。

この出力形態選択部24は制御部2からのメディア選択要求信号を受けて起動され、どのメディアを通じてデータ出力するかを選択するものである。つまり文字コード列や音声、画像等の種々のメディアのうち、どのメディアを通じて情報を通信出力するかを選択するものである。

第65図はこの出力形態選択部24の概略構成図であり、メディア選択制御部24a、入力メディア判定部24b、相手メディア判定部24c、メディア変換テーブル24d、および自己メディア機能テーブル24eを備えて構成される。また第66図はこの出力形態選択部24の概略的な処理の流れを示すものであ

しかし終了選択信号が入力された場合には、上述した如く作成・更新した習熟度表を外部記憶装置の習熟度ファイルに、その該当利用者の識別コードと共に書き込み、これを保存する。そしての一連の処理手続きを終了する。

このようにして作業環境データ収集部25では、システムの操作に関する習熟度のデータを収集しながら、その収集されたデータに従ってその操作を適切にガイダンスするものとなっている。

以上が本ワークステーションの基本的な構成とその機能である。

次に本発明の特徴とするメディア変換方式について説明する。

ここで行われるメディア変換とは、例えば前記音声認識部19にて入力音声を認識処理して文字コード列に変換したり、また前記音声合成部26にて文字コード列を音声合成処理して合成音声を生成したりする処理である。この音声メディアと文字コードメディアとの間の変換のみならず、本ワークステーションは图形認識・图形合成、イメージ

認識・イメージ合成、更には音声やイメージ等の圧縮・伸長処理等の種々のメディア変換機能を備えている。

メディア選択要求信号が与えられるとメディア選択制御部24aは前記制御部2に対してメディア選択動作に必要な入力メディア情報の提供を要求する。そして入力メディア判定部24bに対してメディア情報検出要求とメディア機能識別要求を発する。

入力メディア判定部24bはメディア検出部24fとメディア識別部24gとによって構成される。そして上記メディア選択制御部24aによる情報要求を受けて制御部2から与えられる入力メディアの情報を検出し、且つその検出メディアの機能を識別判定している。この識別判定によって、例えば種々の情報端末に通じるべく入力された情報のメディアが識別検出されるようになっている。具体的には入力メディア判定部24bは、例えば音声データが与えられた場合、その入力メディアが音声であり、またそのメディアの機能がADPCMである等として識別判定する。

しかし後、メディア選択制御部24aは制御部2に対してそのデータ出力の相手先が自己端末(ワークステーション内)の他の機能ブロックであるか、或いは通信回線等を介して接続される別のワークステーションや通信端末であるかを問合せる。そして別のワークステーションや通信端末に対してデータ出力することが指示されると、メディア選択制御部24aは送信相手局に関する識別情報を制御部2に対して要求する。この要求を受けてデータ出力する相手局に関する情報が相手メディア判定部24cに入力される。

相手メディア判定部24cは、相手局識別部24h、相手局メディア識別部24i、機能識別部24jを備えて構成され、前記メディア選択制御部24aからの識別情報判定要求を受けて作動する。そして相手局に関する識別情報をから、先ず相手局の種別を識別し、相手局のメディアを識別する。そしてその相手局メディアの機能を識別する。

具体的には制御部2から与えられる情報に従って、例えばデータ出力(送信)する通信相手局が

他の電話であると判定すれば良い。更に電話でないと判定された場合には、相手機器は電話以外の他の通信機器であると判定するようすれば良い。このような判定処理によって、通信相手局の種別が判定され、その通信相手局が処理可能なメディアの種別が判定される。

しかして通信相手局のメディアが識別判定されると、次にメディア選択制御部24aは、例えば第68図に示すように構成されたメディア変換テーブル24dを参照して、入力メディア、入力機能、相手機器、相手機器メディア、相手機器の機能に対応したメディア変換選択情報を得る。つまり入力情報を通は相手局に送信する場合、該入力情報をどのようにメディア変換すれば良いか、またそれに必要なメディア変換手段はどれであるかの情報を求める。

例えば情報通はすべき入力メディアが音声で、その機能がADPCMであり、相手機器がG3タイプのFAXである場合には、相手機器が処理可能なメディアがイメージであることから、そのメ

自動FAXであり、その通信メディアがイメージであって、その機能がG3タイプである等を識別する。尚、この相手局の識別を、相手局からそのネゴシエーション(ハンドシェーク)機能を用いて直接送られてくる情報に基いて行うようにしても良い。また通信相手局にネゴシエーション機能がない場合には、そのメディア検出機能を機能識別部24jに専用させておけば良い。このようにすれば相手側からのメディア情報信号に従ってその機能識別を行なうことが可能となる。

第67図はこの相手局の識別処理手続の流れを示すものである。この流れに示されるように識別処理は、例えば先ず通信相手局が電話端末か否かを判定し、電話端末である場合には該電話端末からFAX信号が到来するか否かを判定して行われる。

そして相手局が電話であり、FAX信号が到来する場合には、これを相手機器がFAXであると識別すれば良い。また電話であると判定され、FAX信号が到来しない場合には、相手機器は通

ディア変換として

(音声) \rightarrow (コード文字)

(コード文字) \rightarrow (イメージ)

を実行すれば良いことを求める。同時にその変換機能が、

(ADPCM; 音声) \rightarrow (G3; FAX)

によって実現できることを求める。この際、従属的なメディア変換情報が存在すれば、これも同時に求める。

このようにして求められたメディア変換情報が制御部2に与えられ、前記データ出力の形式が選択的に指定される。この結果、前記ADPCMのデータ形式で入力される音声データは、前記音声認識部19にて一旦文字コード列に変換され、更にイメージ合成部27にて文字画像のイメージに変換される。そして通信相手局がFAXであることからG3モードの信号(通信メディア)に変換されて通信装置12、13から送信出力される。このようなメディア変換機能の選択制御によってイメージデータの出力機能を備えた通信相手端末に対して

音声入力された情報が通信出力されることになる。尚、データ出力が自己のワークステーション内部に対して行われる場合には、メディア選択制御部24aは自己メディア機能テーブル24bを参照して、データ出力が可能な出力形式を求める。この情報に従ってメディア選択制御部24aは前記メディア変換テーブル24dの自己メディア変換テーブルを参照し、同様にしてメディア変換情報を求め、これを制御部2に与える。

このようにして求められるメディア変換情報に従って、例えば前述した音声合成部26を用いて文字コードの系列で与えられる文章情報を音声情報にメディア変換したり、或いは音声認識部18を用いて音声情報を文字コード系列の情報にメディア変換される。

尚、通信相手端末が複数種類のメディアに対応可能な場合には、その中の1つを選択してメディア変換を選択制御するようにすれば良い。この際、選択した相手メディアが使用中であれば空きメディアを探し出して選択制御するようにすれば良い。

的にお情報通信することができる。従って種々の形態の端末に対して、その端末が扱い得るメディアで情報を通信出力することが可能となり、通信範囲の拡大と、ワークステーション利用者への便宜の拡大を図ることが可能となる。また種々のメディアを有効に利用した情報通信を行なうことが可能となる等の実用上多大なる効果が認められる。

4. 図面の簡単な説明

図は本発明の一実施例を示すもので、

第1図は本発明に係るワークステーションの特徴的な処理機能を示す図、第2図はワークステーションの概略構成図。

第3図はワークステーションに付随するICカードの外観図、第4図はICカードの構造を示す分解斜視図、第5図はICカードのプリント基板部の構造を示す図、第6図はICカードの半導体集積回路部の構成を示す図。

第7図はワークステーションにおける暗号化処理部の構成を示す図、第8図は暗号・復号化の概念を示す図、第9図は暗号化部の構成図、第10

図には通信相手端末がワークステーションに対して、該ワークステーションが対応可能なメディアの種別の問合せを行なってきた場合には、前記自己メディア機能テーブル24bを参照してその問合せに応答し、該ワークステーションの種別や該ワークステーションが対応可能なメディアの種別を通知するようにしても良い。その上で、通信相手端末との協議によって通信メディアを選択するようにしても良い。

また通信相手端末がメディア選択の自動モードに設定されているか、或いは手動モードに設定されているかを判定してメディア変換の選択制御を行なうようにしても良い。

【発明の効果】

以上説明したように本発明によれば、通信すべき情報の入力メディアを検出すると共に、通信相手端末が処理可能なメディアを検出し、その検出結果に従ってメディア変換の形態を選択するので、通信端末の機能がワークステーションと同一でない場合であっても、その通信相手端末に効果

図は復号化部の構成図、第11図はRSA処理部の構成図。

第12図はワークステーションにおけるイメージ処理部の構成を示す図、第13図はイメージ処理される顔の例を示す図、第14図はイメージデータの構造を示す図。

第15図はワークステーションにおける音声認識部の構成を示す図、第16図は入力音声パターンの例を示す図、第17図は子音の音響的特徴を示す図、第18図は通話ネットワークの例を示す図、第19図は音声認識処理の手続きを示す図、第20図は入力音声に対する部分区間検出を説明する図、第21図は音声認識辞書の学習処理手続きを示す図。

第22図はワークステーションにおける文字認識部の第1の文字認識ブロックの構成を示す図、第23図は認識対象となる文字が記載されるFAX送信原稿用紙の例を示す図、第24図は認識対象文字の切出し処理を説明する図、第25図は文字認識部における第2の文字認識ブロ

クの構成を示す図、

第26図はワークステーションにおける图形認識部の構成を示す図、第27図乃至第30図は图形認識処理を説明する為の図、

第31図はワークステーションにおけるイメージ認識部の構成を示す図、第32図はコード変換装置の構成図、第33図は入力イメージに対する処理例を示す図、第34図はセグメントにおける特徴点検出を示す図、

第35図はワークステーションにおける音声照合部の構成を示す図、第36図はフィルタバンクの帯域分割例を示す図、第37図はフィルタ特性を示す図、

第38図はワークステーションにおける音声合成部の構成を示す図、第39図は規則合成パラメータ生成装置の構成図、第40図は音声パラメータの変換構造を示す図、第41図は音声合成器の構成図、

第42図はワークステーションにおけるイメージ合成部の構成を示す図、第43図および第44

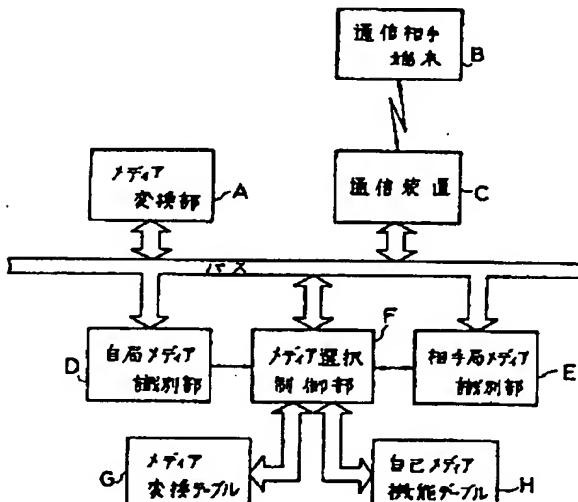
図はイメージ合成処理の概念を示す図、

第45図はワークステーションにおけるデータベース部の構成を示す図、第46図はデータベースのデータ構造を示す図、第47図はリレーションの例を示す図、第48図はリレーションの構造を示す図、

第49図はコマンド対応テーブルの構造を示す図、第50図はワークステーションにおける作業環境データ収集部の構成を示す図、第51図乃至第54図はコマンド部の処理を説明する為の図、第55図はシステム背熟度のデータ収集処理の流れを示す図、第56図は背熟度表の構造を示す図、第57図乃至第64図は作業環境データ収集部の処理を説明する為の図、

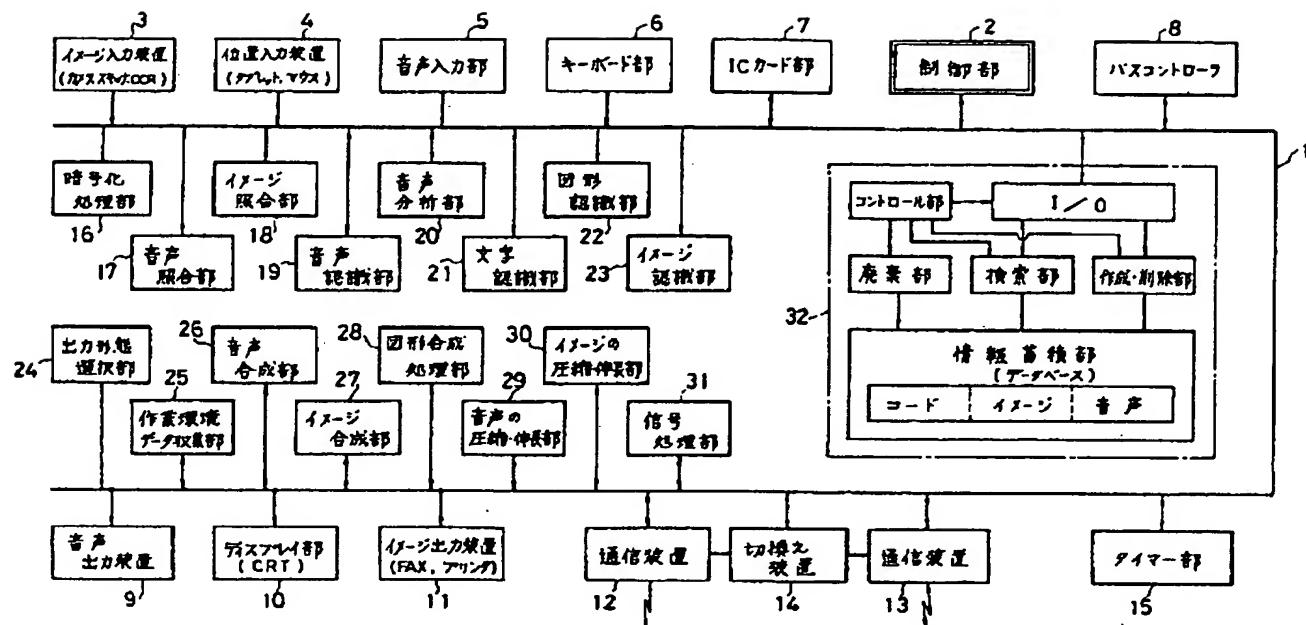
第65図は本発明に係るメディア交換の選択制御を司る出力形態選択部の構成を示す図、第66図は出力形態選択手続きの流れを示す図、第67図は相手局識別処理手続きの流れを示す図、第68図はメディア交換テーブルの構造を示す図である。

1…バス、2…制御部、3…イメージ入力装置、
4…位置入力装置、5…音声入力部、6…キーボード部、7…I Cカード部、8…バスコントローラ、9…音声出力装置、10…ディスプレイ部、
11…イメージ出力装置、12、13…通信装置、14…切換装置、15…タイマー部、16…暗号化処理部、
17…音声照合部、18…イメージ組合部、19…音声認識部、20…音声分析部、21…文字認識部、22…图形認識部、23…イメージ認識部、24…出力形態選択部、25…作業環境データ収集部、26…音声合成部、27…イメージ合成部、28…图形合成部、29…音声の圧縮・伸長部、30…イメージの圧縮・伸長部、31…信号処理部、32…データベース部。

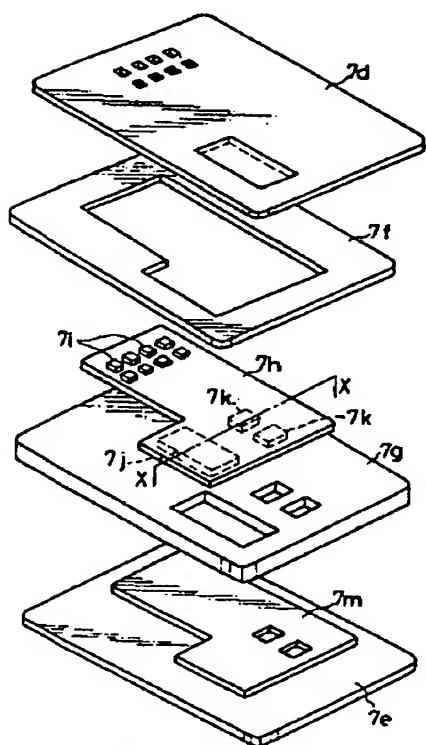


第1図

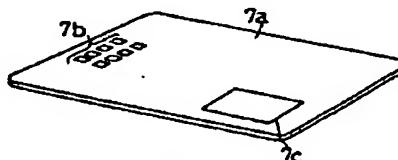
出願人代理人弁理士鈴江武彦



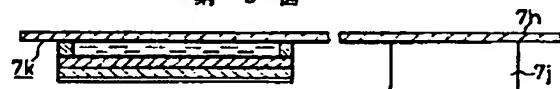
第 2 図



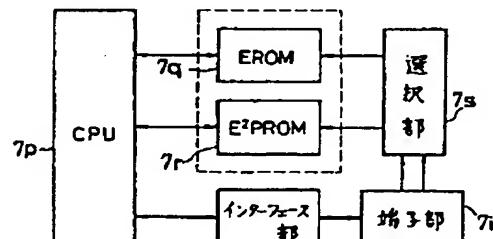
第 4 図



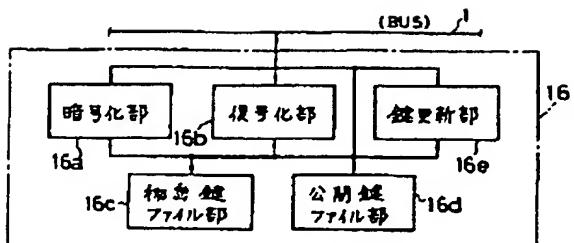
第 3 図



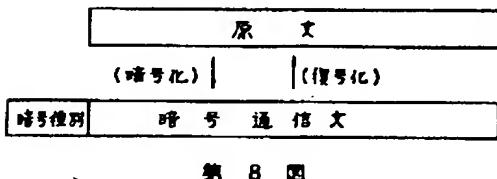
第 5 図



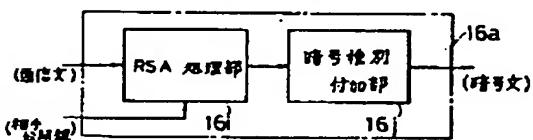
第 6 図



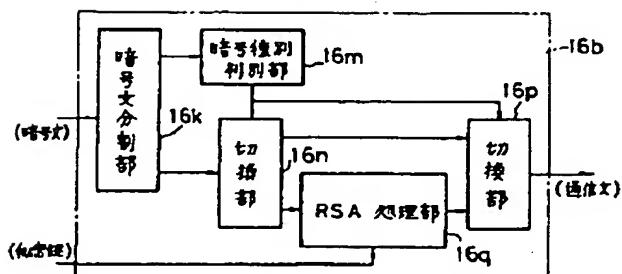
第 7 圖



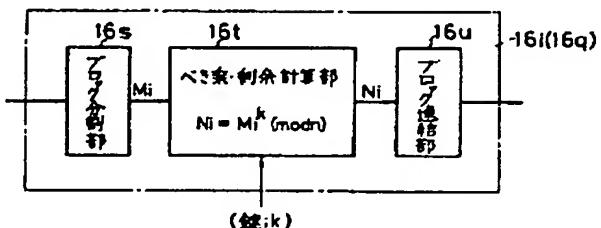
第 8 四



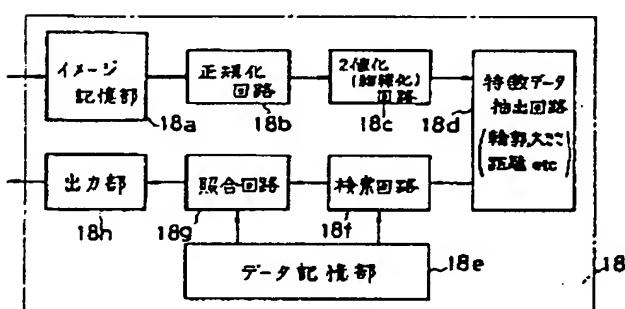
第 9 四



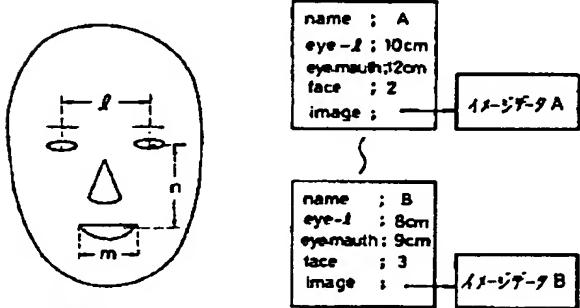
第 10 図



卷一百一

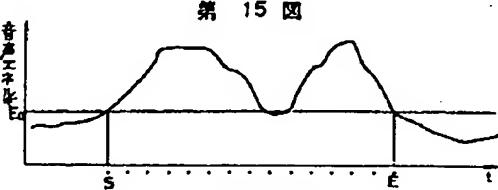


第 12 四

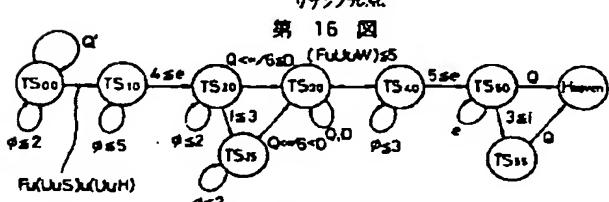


第 13 圖

第 14 国



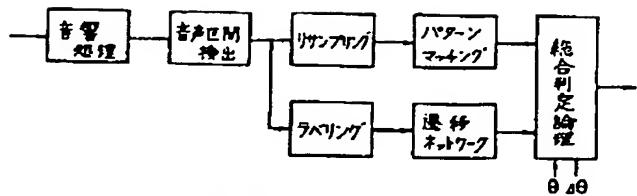
第 15 四



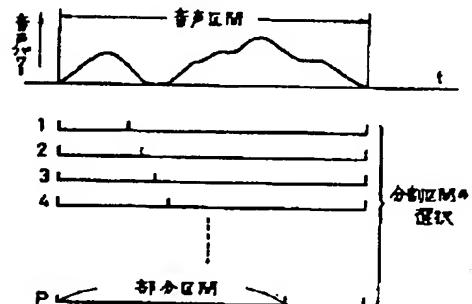
三 18 四

記号	音習的特徴	グロスフィルタの出力の特徴				
		1	2	3	4	(a)
Q	無音性			$E < E_0$		
U	無声性					
F	摩擦性					
W	摩擦性					
B	バズ					
G	附頸性					
N	鼻音性					
S	共鳴性					
H	共鳴性					
M	共鳴性					
D	エキルギー					
V	有声性					(U u F u W)

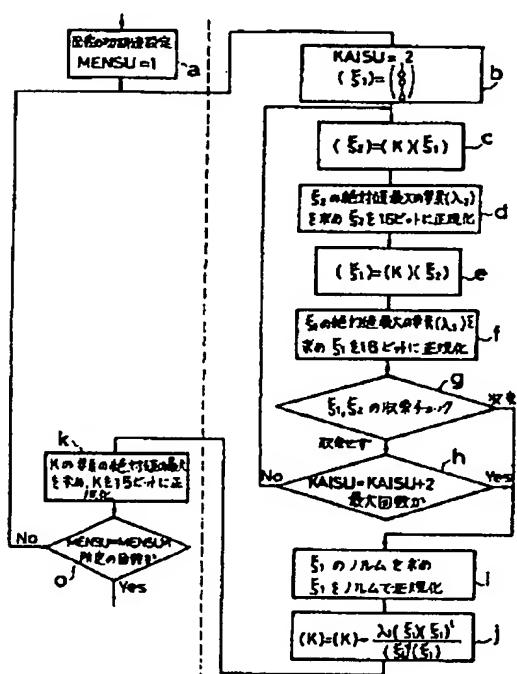
第 17 圖



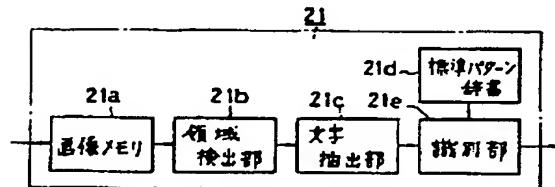
第 19 圖



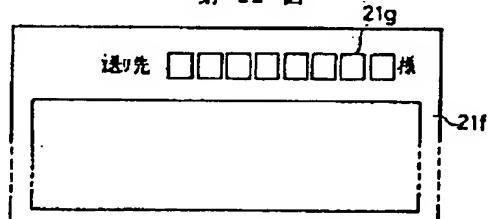
第 20 四



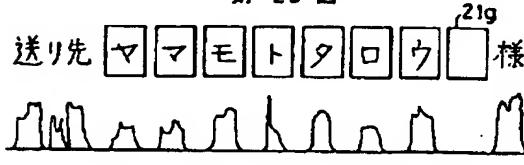
第 21 回



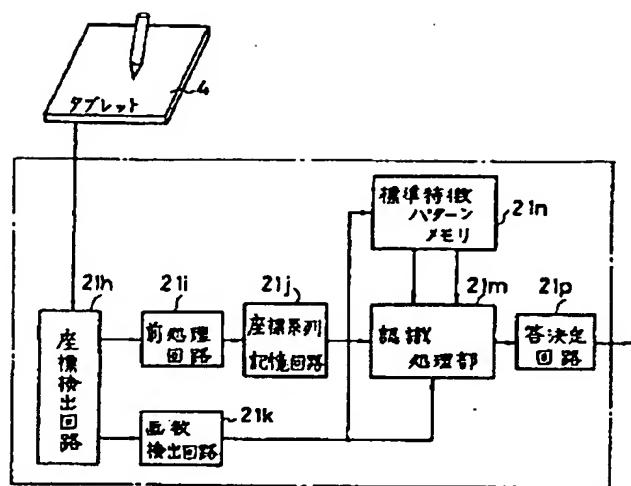
第 22 図



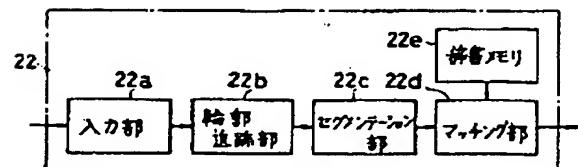
第 23 図



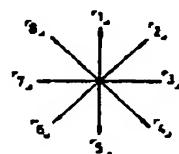
第 24 図



第 25 図



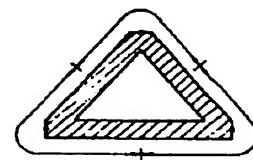
第 26 図



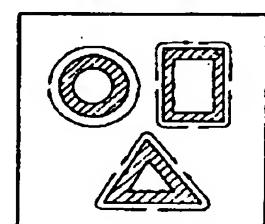
第 27 図



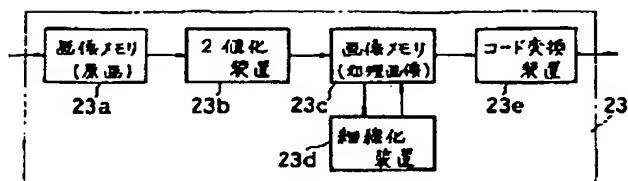
第 28 図



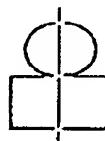
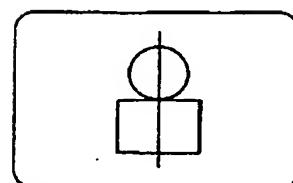
第 29 図



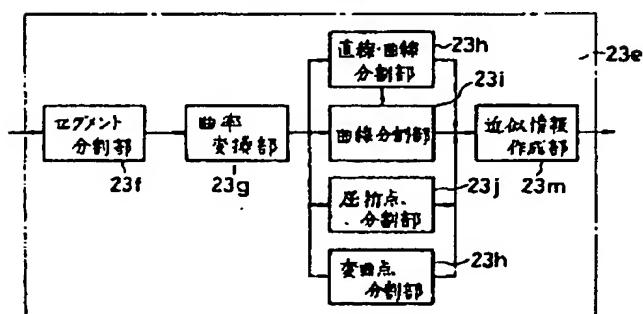
第 30 図



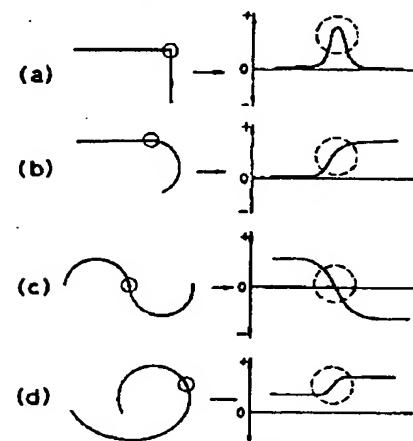
第 31 図



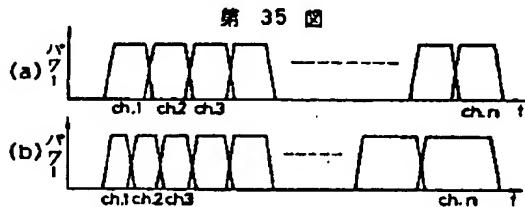
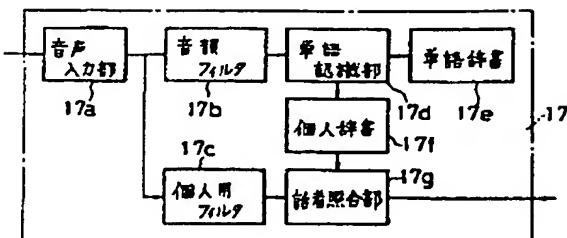
第 33 図



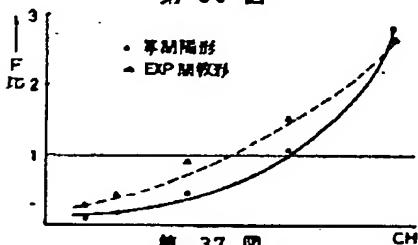
第 32 図



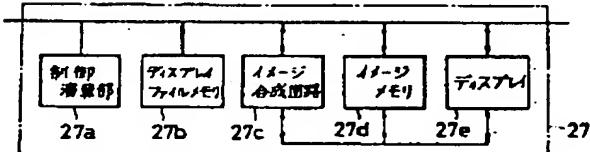
第 34 図



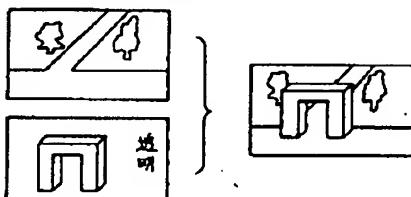
第35図



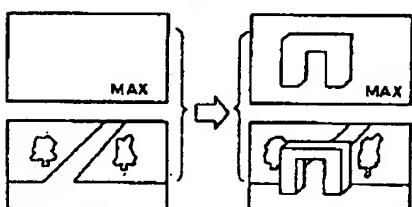
第37図



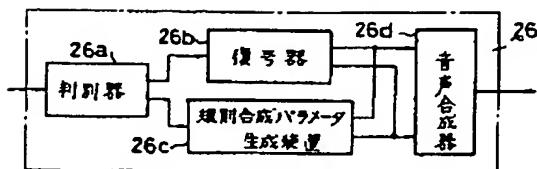
第42図



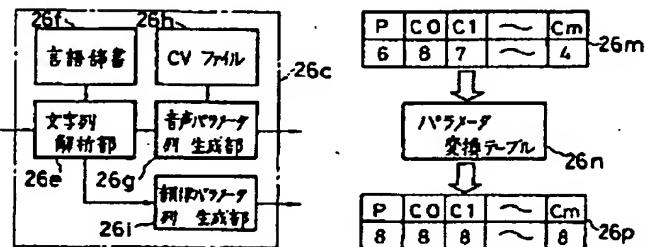
第43図



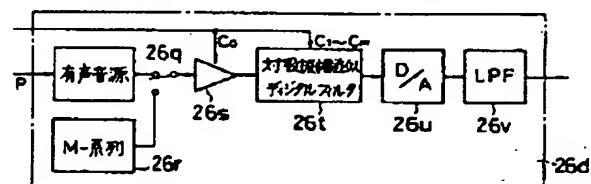
第44図



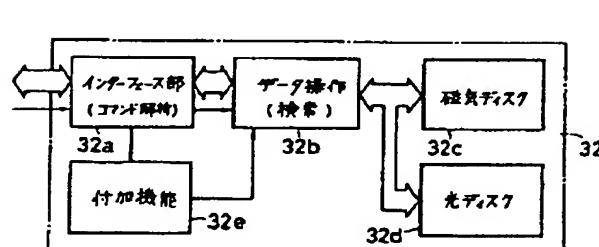
第38図



第39図



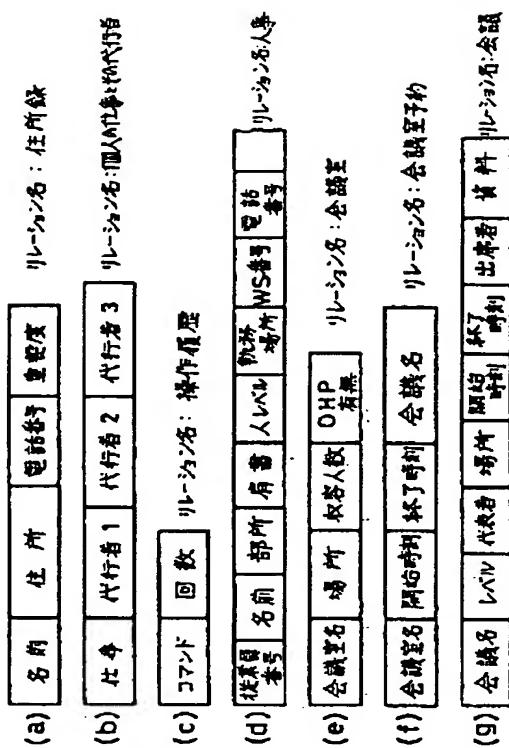
第40図



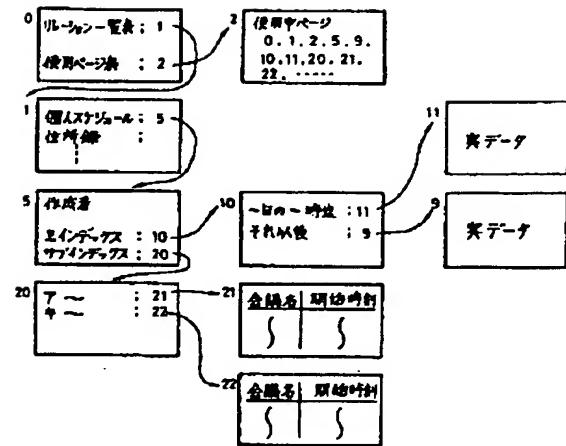
第41図

リレーション名		個人スケジュール						
作成者		△△△△						
読み出し		全員						
追加		○○○○: 所属(技術部)						
変更		人レベル ≥ 5						
データ保蔵情報		開始時刻	終了時刻	種類	名前	場所	行数	設定者
データ		86.7.7 9:00	86.7.7 10:30	会議	会議室	A会議室	4	○○○
		○	○	○	○	○	○	○

第46図



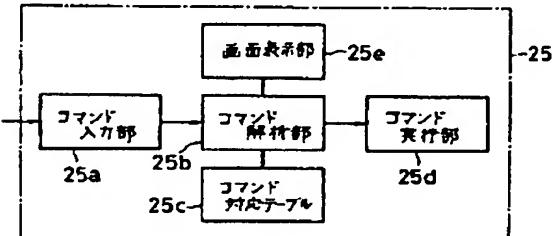
第 47 図



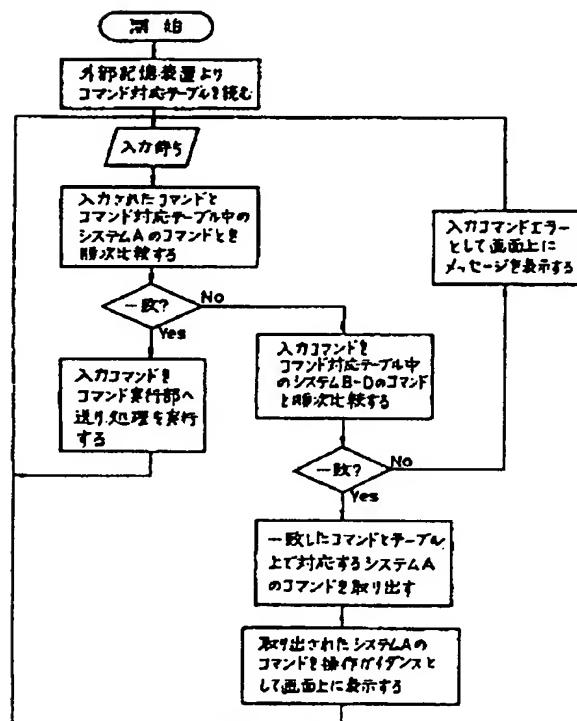
第 48 図

システム A	システム B	システム C	システム D	-----
DELETE	DEL	ERASE	REMOVE	-----
CATA	DIR	CLIST	CATLIST	-----
CREATE	CRE	MAKEFL	NEWFILE	-----
EDIT	ED	EDITFL	FILEED	-----
COPY	COPY	COPY	FILECOPY	-----
...

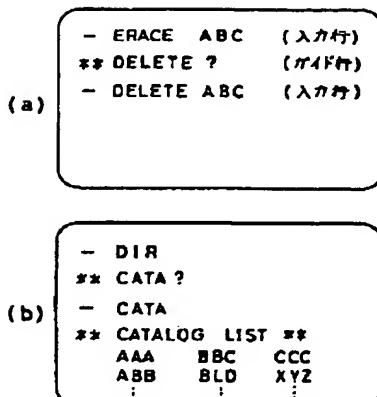
第 49 図



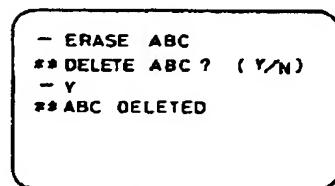
第 50 図



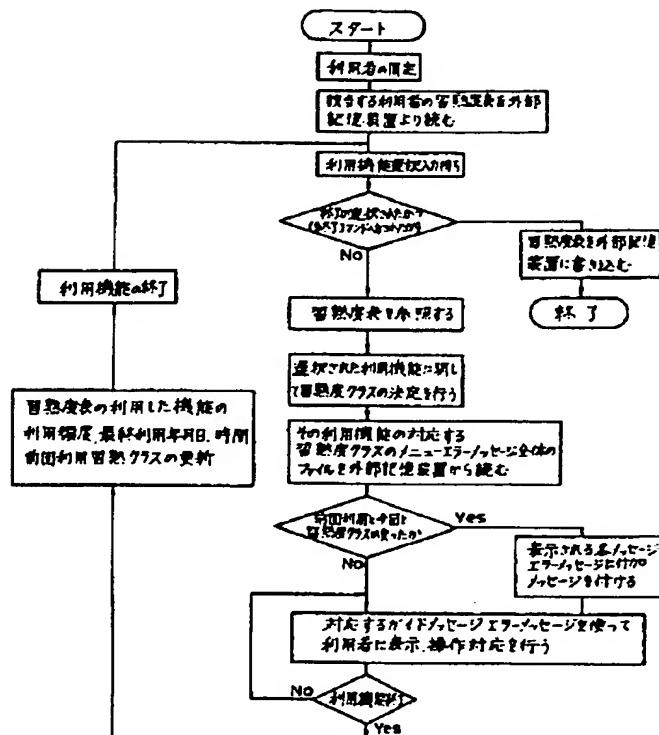
第 51 図



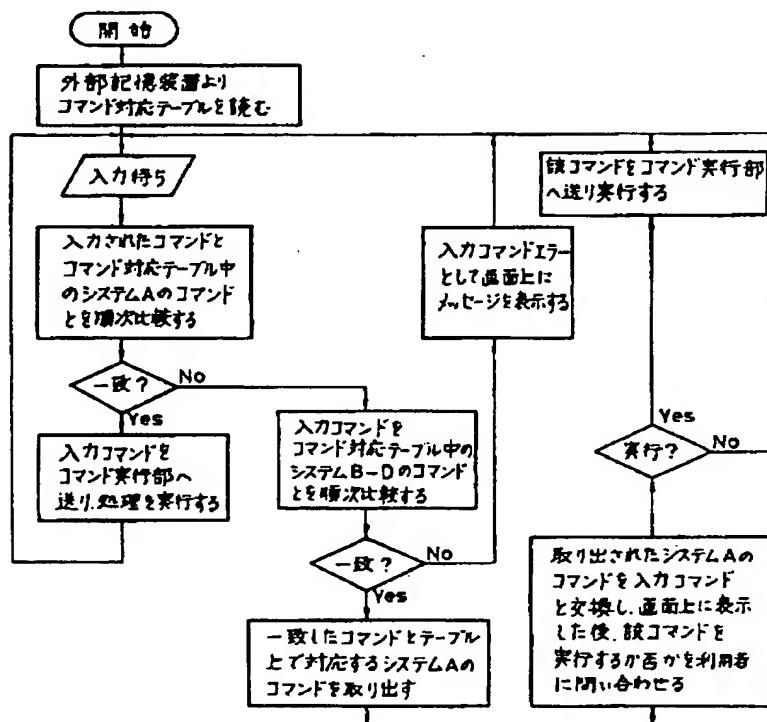
第 52 四



第 54 四



第 55 四



第 53 圖

利用機器名	利用頻度	最終利用日時	利用者登録ID	利用登録ID	登録日
エディタ 基本操作	98	1983/12/15/07	日	B	1
エディタ 高級操作	---	---	A	A	2
ワードプロセッサー フォーマット	---	---	A	A	1
日本語文書作成システム 基本操作	---	---	B	---	---
日本語文書作成システム 高級操作	---	---	A	---	---
リストアインターフェース	---	---	A	---	---
リストエンパイラ	---	---	A	---	---
ファイルアライアライ	---	---	A	---	---
電子メール	---	---	C	---	---
電子ファイル	---	---	B	---	---
OA Cafe	---	---	---	---	---
文献検索システム	---	---	---	---	---

第 56 図



第 57 図

装置登録履歴		付: カロナーベージ
1	A - B	*あなたは 十分正確だったので階層メニューに登りました。
2	B - C	*あなたは 許可されたので追加登録が使えるメニューに登りました。
3	C - B	*あなたは 最近利用されていませんのでサブファイルメニューになりました。
4	B - A	*あなたは 最近利用されていないので、便利やすいメニューになっています。

第 58 図

** 作成したデータを
ファイルに格納します。**

データを入れる場所の名前を入れて下さい。
まず装置名です。

装置名

装置名とは、あなたの日常使う機器装置
の名前です。例とば [DK00]で表します。
押さなければ、キーボードのキー[F14]を押せば
あなたの使っている装置名の候補が表示され
から選択して下さい。

・装置とは... フォルダ名
・ファイルとは... フォルダ
・サブファイルとは... フォルダ
サブファイル

第 59 図

** 入力の方法が判らないようですね。 キー[F14]を押すと
装置名の候補が表示されます。 修正したい場合は、入力前に修正キー
を押して下さい。 なお キーボードのキー[F15]を押すとまとめて便利です。

第 60 図

** 作成したデータを
ファイルに格納する**

装置名

ファイル名

サブファイル名

有効期限

装置名は例とば [DK00]のように英文字と
数字の2桁づつ表します。
なお、あなたの使っている装置ファイル名が
選択する場合は、キー[F15]を押して
選択して下さい。

第 61 図

** 作成したデータの格納 **

装置名 ファイル名 サブファイル名 有効期限

なおデフォルトの値を利用されるのでしたら、あ
らかじめ、[F14]キーでセットしておけば
以後自動的に格納されます。この機能
は便利です。

あなたのサブファイルの利用情報は、次
のようになります。あと…% 使用すると
ファイルがオーバーします。

サブファイル名

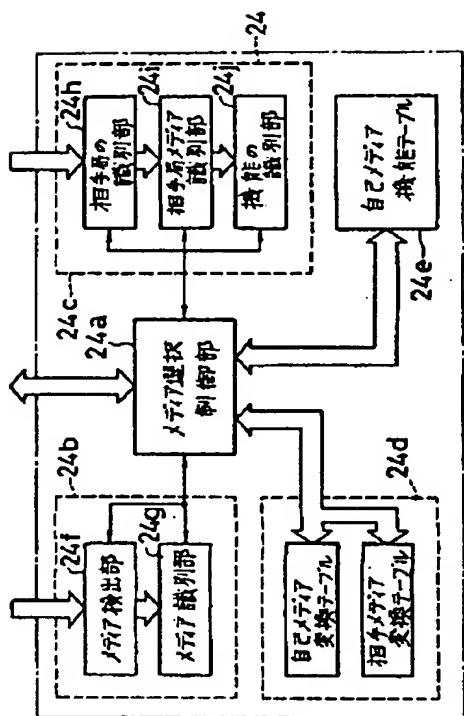
第 63 図

** 装置名の入力が誤っています。正しい装置名を入れて下さい。

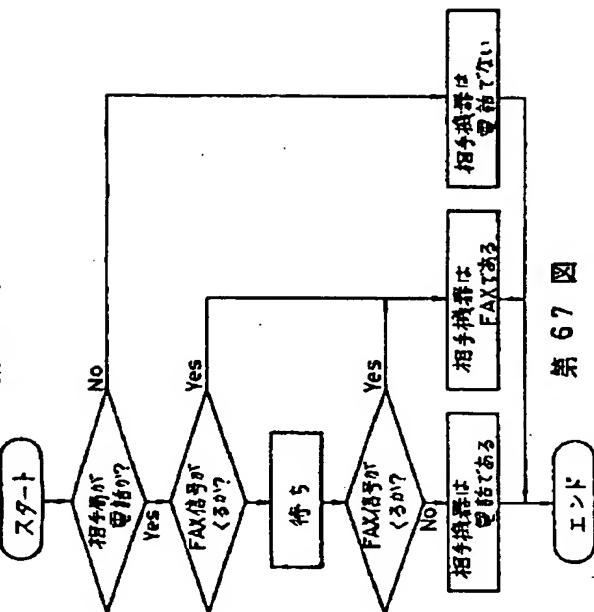
第 62 図

** ファイル名記入誤り

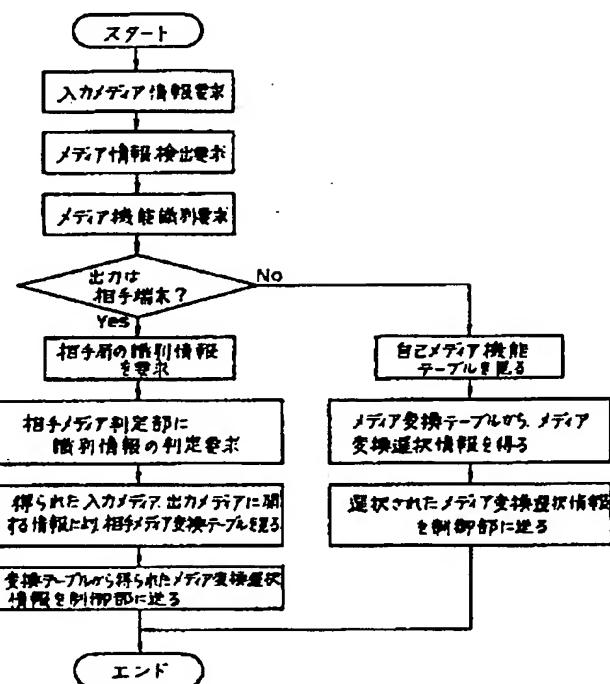
第 64 図



第 65 図



第 67 図



第 66 図

入力 ノイズ	音 声	音 声	コード文字	マルチピット	音 声
入力端子 ADPCM	ADPCM	基本PCM	レコード-1724 (送信端子723)	Log PCM	
相手端子 FAX	FAX	FAX	電話	マルチピット	
相手端子 FAX	イメージ	イメージ	音 声	イメージ	ミクスマード (送信端子5)
相手端子 の端子 G III	G III	G IV	Log PCM	G III	音 声-ADPCM イメージ
相手端子 の端子 G III	音 声-コード文字 G III-イメージ				
送信端子 ADPCM-GIII	ADPCM	2441-F -2441-F	Log PCM	レコード-F -GIII	Log PCM -ADPCM
送信端子 ADPCM	—	音 声-コード文字 G III-イメージ	—	—	—
送信端子 ADPCM	—	音 声-コード文字 G III-イメージ	—	—	—

第 68 図

第1頁の読み

⑥Int.Cl.
H 04 L 13/00

識別記号 305 厅内整理番号 B-7240-5K

⑦発明者 根岸 殿 神奈川県川崎市幸区小向東芝町1番地 株式会社東芝総合
研究所内